

GenCore version 5.1.3
Copyright (c) 1993 : 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2002, 15:04:13 : Search time 4357 Seconds
(without alignments)
11341.878 Million cell updates/sec

Title: HSYBL_COPY_453_2150
Perfect score: 1698
Sequence: 1 AGAACCATATGTCGGCCAA.....CACCCCGGAGGAGCGCCAG 1698

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:**

1: gb_ba:**
2: gb_htg:**
3: gb_in:**
4: gb_om:**
5: gb_ov:**
6: gb_pat:**
7: gb_ph:**
8: gb_pl:**
9: gb_pr:**
10: gb_ro:**
11: gb_sts:**
12: gb_sy:**
13: gb_un:**
14: gb_vi:**
15: em_ba:**
16: em_fun:**
17: em_hum:**
18: em_in:**
19: em_mu:**
20: em_om:**
21: em_or:**
22: em_ov:**
23: em_pat:**
24: em_ph:**
25: em_pl:**
26: em_ro:**
27: em_sts:**
28: em_un:**
29: em_vi:**
30: em_htg_hum:**
31: em_htg_inv:**
32: em_htg_other:**
33: em_htg_mus:**
34: em_htg_pln:**
35: em_htg_rod:**
36: em_htg_man:**
37: em_htg_vrt:**
38: em_sy:**
39: em_htgo_hum:**
40: em_htgo_mus:**
41: em_htgo_other:**

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1698	100.0	2550	9	HSYB1	X96666 H. sapiens Y
2	1672.4	98.5	161855	9	AC098484	AC098484 Homo sapi
3	456.6	26.9	659	9	HUMYZ92E04	AF086102 Homo sapi
C	186.8	11.0	195623	2	AC101988	AC101988 Mus muscu
5	107.4	6.3	1548	6	AX014868	AX014868 Sequence
6	101.4	6.0	144902	2	AC016114	AC016114 Homo sapi
7	101.4	6.0	162144	9	CNS06C85	AL391261 Human chr
C	101.4	6.0	184656	9	CNS05TDB	AL356021 Human chr
C	96.6	5.7	189367	9	AL162733	AL162733 Human DNA
10	95	5.6	99886	9	AL135841	AL135841 Human DNA
11	89.4	5.3	395	6	AX321258	AX321258 Sequence
12	89	5.2	1481	6	AR117695	AR117695 Sequence
13	89	5.2	1481	6	AX409453	AX409453 Sequence
14	89	5.2	1481	9	HUMYB1A	J03827 Y box bindi
15	87.4	5.1	2168	9	HUMPSDBPB	L37516 Homo sapien
16	86.6	5.1	162996	2	AC051655	AC051655 Homo sapi
17	85.2	5.0	410	6	AX321869	AX321869 Sequence
18	85.2	5.0	926	9	BC000064	BC000064 Homo sapi
19	83	4.9	3073	6	AR083654	AR083654 Sequence
20	83	4.9	3073	6	I22491	I22491 Sequence 31
21	81.4	4.8	1468	9	HUMNSEP	M82234 Human nucle
22	81.4	4.8	1474	9	HUMRNABP	M85234 Human nucle
23	80.6	4.7	142184	2	AC037428	AC037428 Homo sapi
24	80.6	4.7	143604	9	AC093610	AC093610 Homo sapi
25	80.6	4.7	310021	2	AC092883	AC092883 Homo sapi
26	79	4.7	1552	9	BC002411	BC002411 Homo sapi
C	79	4.7	77948	2	AC018982	AC018982 Homo sapi
28	78.8	4.6	381	6	AX365822	AX365822 Sequence
29	77.8	4.6	241390	2	AC093965	AC093965 Rattus no
30	77.4	4.6	545	6	AX321285	AX321285 Sequence
31	77	4.5	1554	9	BC010430	BC010430 Homo sapi
32	76.8	4.5	142565	10	AL592545	AL592545 Mouse DNA
33	76.8	4.5	197658	2	AL646094	AL646094 Mus muscu
34	76.6	4.5	7218	6	I66494	I66494 Sequence 14
35	76.6	4.5	179402	9	AC079328	AC079328 Homo sapi
C	76.6	4.5	187005	2	AC024309	AC024309 Homo sapi
37	75.4	4.4	1452	9	HUMDBPB	M24070 Human DNA-b
38	74	4.4	307	9	HSBFB31	Z70765 H. sapiens m
39	73.8	4.3	1503	4	OCU16821	OCU16821 Oryctolagus
C	72.8	4.3	169224	2	AC113562	AC113562 Sus scrofa
C	72.6	4.3	159409	2	AC092900	AC092900 Homo sapi
42	72.6	4.3	160877	2	AC026560	AC026560 Homo sapi
C	72.6	4.3	186333	2	AC027493	AC027493 Homo sapi
C	72.6	4.3	187215	9	AC092915	AC092915 Homo sapi
45	67.4	4.0	1407	10	BC013450	BC013450 Mus muscu

ALIGNMENTS

RESULT 1

HSYB1
LOCUS HSYB1 2550 bp DNA linear PRI 08-JUL-1996
DEFINITION H. sapiens YB-1 gene promoter region.
ACCESSION X96666
VERSION X96666.1 GI:1403348
KEYWORDS promoter region; Y box binding protein; YB-1 gene.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2550)
AUTHORS Makino, Y., Ohga, T., Toh, S., Koike, K., Okumura, K., Wada, M.,
Kuwano, M. and Kohno, K.
TITLE Structural and functional analysis of the human Y-box binding

	Query Match	100.0%;	Score 1698;	DB 9;	Length 2550:	
	Best Local Similarity	100.0%;	Pred. No. 0;			
	Matches 1698;	Conservative	0;	Mismatches	0;	Gaps
QY	1	AGAACCATATGTCGGCCAAAGCTTATATTAACTATTGGCTCTCACAAAAAAGTTT	60			
Ddb	453	AGAACCATATGTCGGCCAAAGCTTATATTAACTATTGGCTCTCACAAAAAAGTTT	512			
QY	61	CCTACCCCTTGCTTAAGCAATTTGGTGATGCCCTACAGTAGTCACACATAGGAATGAA	120			
Ddb	513	CCTACCCCTTGCTTAAGCAATTTGGTGATGCCCTACAGTAGTCACACTAGGAATGAA	572			
QY	121	AGACTCTGAATTGCATTTAGAGGGTTTAGGACTCAGTCAAGACGCCTTTAGGAGGTG	180			
Ddb	573	AGACTCTGAATTGCATTTAGAGGGTTTAGGACTCAGTCAAGACGCCTTTAGGAGGTG	632			
QY	181	GAACCTCACAGGCCTAGACGGCATTGTTAGGGGTAAAGGAATACATGCAATTAGCCCTCCCTAACT	240			
Ddb	633	GAACCTCACAGGCCTAGACGGCATTGTTAGGGGTAAAGGAATACATGCAATTAGCCCTCCCTAACT	692			
QY	241	GCTTTCTCTACTTCGATTTCCTTGCCTCTCGAACCCGANTTCTCCACTCCGACGCAATTTT	300			

[illegible]

3398	3599	3657	3675	11517	11316
2503	2527	14158	13689	1044	1024
3341	3393	2386	2336	5448	5071
5997	5970	4939	5026	2779	2792
11401	11247	2187	2206	3013	3120
1470	1443	6457	6489	3869	3810
12564	12727	1771	1836	9708	9870
6660	6745	3788	3991	12923	12794
4555	4619	4838	4792	4563	4622
1729	1734	3160	3127	287	<800
406	<800	4088	4313	9519	9493
2213	2193	5832	6201	2229	2229
6922	6981	942	971	2610	2628
4531	4619	4341	4619	5661	5695
1381	1370	9735	9702	576	<800
1386	1370	4886	4792	820	830
1168	1168	15773	15940	3668	3654
15253	15479	398	<800	9025	8972
3613	3599	5122	5509	5092	5071
8174	8090	1336	1324	1026	1024
2074	2086	10995	10821		
714	<800	4329	4313		
6275	6239	16	<800		
3627	3599	1338	1324		
		978	971		
		2203	2206		
		2343	2336		
		489	<800		
		6300	6489		
		376	<800		
		2981	2969		
Location/Qualifiers					
1. .161835					
/organism="Homo sapiens"					
/db_xref="taxon:9606"					
/chromosome="1"					
/clone_lib="RP5-994D16"					
/clone_lib="RPCI human PAC library 5"					
1. .4					
/note="Single subclone region"					
161830. .161835					
/note="Single subclone region"					
misc_feature					
misc_feature					

BASE COUNT	43236	a	36046	c	37167	g	45386	t	
ORIGIN									
Query Match	98.5%; Score 1672.4; DB 9; Length 161835;								
Best Local Similarity	99.8%; Pred. No. 0;								
Matches 1695; Conservative	0; Mismatches 1; Indels 2; Gaps 2;								
QY	1	AGAACCATATGGTCGCGCAAGCTTATATTA	ACTATTTGGCTCTCACAGAAAAAGTTT	60					
Db	28812	AGRACCATATGGTCGCGCAAGCTTATATTA	ACTATTTGGCTCTCACAGAAAAAGTTT	28871					
QY	61	CCTACCCCTTGCTCTAAGCAATTTGGTGAT	GATGGCTTACAGTAGTACAGTAGGAATGAA	120					
Db	28872	CCTACCCCTTGCTCTAAGCAATTTGGTGAT	GATGGCTTACAGTAGTACAGTAGGAATGAA	28931					
QY	121	AAGACTCTGAATTCACATTTAGAGGGTTT	AGGACTGACTCAAGAACCCCTTTAGGAGGTG	180					
Db	28932	AAGACTCTGAATTCACATTTAGAGGGTTT	AGGACTGACTCAAGAACCCCTTTAGGAGGTG	28991					
QY	181	GAACCTCACAGGCTTAGACGGCATTGGT	AGGGTAAAGCAATACGCTTACGCTCCCTAACT	240					
Db	28992	GAACCTCACAGGCTTAGACGGCATTGGT	AGGGTAAAGCAATACGCTTACGCTCCCTAACT	29051					
QY	241	GCTTTCTCTACTTCCATTCCTTCCCTCT	CGCAACCCATTCTCCACTCCCGCAGCCATTTT	300					
Db	29052	GCTTTCTCTACTTCCATTCCTTCCCTCT	CGCAACCCATTCTCCACTCCCGCAGCCATTTT	29111					
QY	301	TAAAAAGATGCCCTCCCTACTTATGAC	TCTAAATTTGCTTCTCACTTCCCTCAG	360					
Db	29112	TAAAAAGATGCCCTCCCTACTTATGAC	TCTAAATTTGCTTCTCACTTCCCTCAG	29171					
QY	361	GATATATTTCCAAATTAATATACCTA	AGTGACTGACCTGCAACCCCTCTGCAACCC	420					
Db	29172	GATATATTTCCAAATTAATATACCTA	AGTGACTGACCTGCAACCCCTCTGCAACCC	29231					
QY	421	TCGAGTCTTACTGAATCTTACTGCTC	ATTCCCGAGATCTACCTCTTCTCGCCCTG	480					
Db	29232	TCGAGTCTTACTGAATCTTACTGCTC	ATTCCCGAGATCTACCTCTTCTCGCCCTG	29291					
QY	481	CCTGTGCGGGAAGTCAGCCCTCCAC	CTTCCCTTCCCTTCCCTTCCCTTCCCT	540					
Db	29292	CCTGTGCGGGAAGTCAGCCCTCCAC	CTTCCCTTCCCTTCCCTTCCCTTCCCT	29351					
QY	541	GGTTTTCAGCTCTGGAGTATTTACCG	TGTTGGCTGTTTAAATTTCTGCTCCAT	600					
Db	29352	GGTTTTCAGCTCTGGAGTATTTACCG	TGTTGGCTGTTTAAATTTCTGCTCCAT	29411					
QY	601	GGCAGAACTGACTCGCGAATCTTCC	ATCCCATCCCGAGCGATAGTAGCGCTT	660					
Db	29412	GGCAGAACTGACTCGCGAATCTTCC	ATCCCATCCCGAGCGATAGTAGCGCTT	29471					
QY	661	CGAAGAAGTGGTGGGAGGACTTCAG	TAAATCAGGTGGCAGCCTCAATTTATCG	720					
Db	29472	CGAAGAAGTGGTGGGAGGACTTCAG	TAAATCAGGTGGCAGCCTCAATTTATCG	29530					
QY	721	TCTGAAACGTGGATAGTAATCCCT	CTCTATCACGTGGTGTTCAGGAATAA	780					
Db	29531	TCTGAAACGTGGATAGTAATCCCT	CTCTATCACGTGGTGTTCAGGAATAA	29590					
QY	781	CAAAACAGGCTAGCTTGTTCATAA	TAATGTGAGTTGAATTTTGGTGCAGTA	840					
Db	29591	CAAAACAGGCTAGCTTGTTCATAA	TAATGTGAGTTGAATTTTGGTGCAGTA	29650					
QY	841	GAAGAAGATGTAATACCTTGGAA	AGGAGACACACATTTTTTAAATATAT	900					
Db	29651	GAAGAAGATGTAATACCTTGGAA	AGGAGACACACATTTTTTAAATATAT	29710					
QY	901	CGGATCAGAAGGAGGTCCTCCCAT	GGAGCACACCTCGCCCTAAACATGCT	960					
Db	29711	CGGATCAGAAGGAGGTCCTCCCAT	GGAGCACACCTCGCCCTAAACATGCT	29770					
QY	961	TCCATAGCCTCGGTGGTCCCTCC	CAAGGTGACTGCTCCGACAAAAAGG	1020					

Qy	1392	GGCGCGCGGAGCGAGGAAACGGTTGTAGTCTGACTGAATTAGCGCCCAAGGT-CCAA	1450
Db	168481	GG----CGGCGGAGCCCTGAGAAGCTTGTGGTCGGCGGAATAGTCTAAAGATACCAA	168426
Qy	1451	TGAGAATTGGAGGACTGATAAAATATTAGCCAAATAGAGCTTAGGGATTGGGTGAGGTGGG	1510
Db	168425	TGAGAGCGCAGAGCCGCGGAGCGCTGAGCCAAATAGGCGCTAGAGA-CCCGGTCAAGTGGT	168367
Qy	1511	CAGATTGACACTACACTGTGGCGACTGAAACCCCTAGGCGGGGTGCTGCTGATGGCGCTTA	1570

QY 1311 CAGATGACAGTACCACCTGGCCAGTGAACAACGCCTAGGGCGGGTCCGTCGTAGGGCTTA 1570

	Dy	168366	GAGATGACAAACGATCGGGACCCTAATGGGGAGGAGCAGGAGCCGGTGTATTTCGCCGGGCCTTA	168307
	QY	1571	TCCGCGCTGTCCC GCCCATTTCTCGCTACTTGCTTAGTCGGTACGGGAGCGGAGCGACGCC	1630
	Dy	168306	TCCGCGCGCTGTCCC GCCCATTTCTCGCTACTTGCTTAGTCGGTACGGGAGCGGAGCGACGCC	168247
	QY	1631	AGAGAGCCCTGAGCAGCCCCACCGCCCGCCGCGCCCTAGTTACCATCACACCCCGGAGG	1690
	Dy	168246	AGAGAGCCCTGAG - AGCCCCACCGCCCGCCGCGCCCTAGTCACCATCACA - CCCGGAGG	168189
	QY	1691	AGCCGCCAG 1698 	
	Dy	168188	AGCCGCCAG 168181	
	RESULT	5		
	LOCUS	AXO14868	Sequence 58 from Patent WO9953040.	PAT 07-SEP-2002
	DEFINITION	Sequence 58 from Patent WO9953040.	linear	
	ACCESSION	AXO14868		
	VERSION	AXO14868.1	GI:10041135	
	KEYWORDS	.		
	SOURCE	human. ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. REFERENCE 1 (bases 1 to 1548) AUTHORS Schmitt,A., Specht,T., Dahl,E., Hinzmänn,B., Rosenthal,A. and Pilarsky,C. TITLE Human nucleic acid sequences from ovarian tumour tissue JOURNAL Patent: WO 995304-A 58 21-OCT-1999; SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN BRND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN (DE); PILARSKY CHRISTIAN (DE)		
	FEATURES	source Location/Qualifiers . . . 1..1548 /organism="Homo sapiens" db_xref="taxon:9606"		
	BASE COUNT	454 A 386 C 403 G 305 T		
	ORIGIN			
	Query Match	6.3%	Score 107.4;	DB 6; Length 1548;
	Best Local Similarity	99.1%;	Pred. No. 1.6e+20;	
	Matches 108;	Conservative	0; Mismatches 1;	Indels 0; Gaps
	Dy	1590	CTCCCTACTTCGATCCGCTACCGCCGCCGAGCGGACGCCGACCGCCGCGCCCTGACCGCCGCC	1649

Db	1	CTCGCTAGTTTCGATCGGTACGGGAGCGGAGCGGACCCAGAGAGCCCTGAGCAGCCC	60
Oy	1650	CACGGCCGCGCCGGCGCTAGTTACCATCACACCCGGGAGGAGCCCGCAG	1698
Db	61	CACGGCCGCGCGCGCTAGTTACCATCACACCCGGGAGGGGCCGCGAG	109
RESULT 6			
AC016114			
LOCUS			
DEFINITION	AC016114	144902 bp	DNA linear HTG 01-APR-2001
	Homo sapiens clone Rpl1-27C17,	WORKING DRAFT SEQUENCE,	8 unordere
	pieces.		
ACCESSION	AC016114		
VERSION	AC016114.3	GI:7382202	
KEYWORDS	HTG; HTGS PHASE1; HTGS DRAFT.		

AC016114	144902 bp	linear	HTG 01-APR-20
LOCUS			
DEFINITION	Homo sapiens clone RP11-27C17, WORKING DRAFT SEQUENCE, 8 unordered		
	pieces.		
ACCESSION	AC016114		
VERSION	AC016114.3	GI:7382202	
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.		
SOURCE	Homo sapiens.		

KEYWORDS HTG. human.
SOURCE ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 189367)
JOURNAL Direct Submission
COMMENT Submitted (16-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Jul 29, 2001 this sequence version replaced gi:14329913.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em., EMBL; Sw.,
SWISSPROT; Tr., TrEMBL; Wp., WormPEP; Information on the WormPEP
database can be found at
http://www.sanger.ac.uk/projects/C.elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr9
Rpl1-397F23 is from the library RPCI-11.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACE3.6
This sequence is the entire insert of clone Rpl1-397F23 The true
left end of clone Rpl1-412H14 is at 176727 in this sequence. The
true right end of clone Rpl1-388N6 is at 96050 in this sequence.

FEATURES
source Location/Qualifiers
1..189367
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/clone="Rpl1-397F23"
/clone_lib="RPCI-11.2"
39854..39882
/note="Single clone region. Sequence from reads from a
short insert library derived from a single pUC clone.
Restriction digest data confirm the assembly."
74553..75378
/note="CpG island"
/evidence="not_experimental"
74855..74911
/note="Single clone region. Sequence from reads from a
short insert library derived from a single pUC clone.
Restriction digest data confirm the assembly."

BASE COUNT 53387 a 41503 c 40875 g 53602 t
ORIGIN

Query Match 5.7%; Score 96.6; DB 9; Length 189367;
Best Local Similarity 96.1%; Pred. No. 1e-16;
Matches 99; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1596 AGTTCGATCGGTAGCGGAGCGGAGCGGACCCAGAGAGCCCTGAGCAGCCCGCCAGCC 1655
Db 111342 AGTTCGATCGGTAGCGGAGCGGAGCGGACCCAGAGAGCCCTGAGCAGCCCGCCAGCC 111283
QY 1656 CCGCCGCCCGCTAGTTACCATCACACCCCGGAGGAGCGCGAG 1698
Db 111282 CACCCGTGGCCCTAGTACCATCACACCCCGGAGGAGCGCGAG 111240

RESULT 10
AL135841/c
LOCUS DEFINITION
Human DNA sequence from clone Rpl1-327L3 on chromosome 9p13.1-13.3.
Contains a putative novel gene, a PGAM1 (phosphoglycerate mutase 1
(brain)) pseudogene, the gene for a novel 7 transmembrane receptor
(rhodopsin family) (olfactory receptor like) protein, an NSEPI
(nuclease sensitive element binding protein 1) (Yb1) pseudogene,
ESTs, STSS and GSSs, complete sequence.
AL135841
AL135841.11 GI:7529194
HTG; NSEPI; olfactory receptor; PGAM1; rhodopsin; Yb1.
human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 99886)
Direct Submission
Submitted (20-APR-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Apr 9, 2000 this sequence version replaced gi:7378508.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., EMBL; Sw., SWISSPROT; Tr., TrEMBL; Wp., WormPEP; Information
on the WormPEP database can be found at
http://www.sanger.ac.uk/projects/C.elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr9
Rpl1-327L3 is from the library RPCI-11.2 constructed at the Roswell
Park Cancer Institute by the group of Pieter de Jong. For further
details see http://bacpac.med.buffalo.edu/
VECTOR: pBACE3.6
IMPORTANT: This sequence is not the entire insert of clone
Rpl1-327L3 it may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true right end of clone Rpl1-327L3 is at 99886 in this
sequence. The true left end of clone Rpl1-113A10 is at 64692 in
this sequence. The true right end of clone Rpl1-112J3 is at 100 in
this sequence.

FEATURES
Location/Qualifiers
1..99886
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/map="p13.1-13.3"
/clone="Rpl1-327L3"
/clone_lib="RPCI-11.2"
4..752
/note="L1M4 repeat: matches 1256. .2036 of consensus"
795..1824
/note="L1M4 repeat: matches 4573. .5641 of consensus"
1824..2482
/note="L1M4 repeat: matches 551. .1226 of consensus"
2469..2811
/note="L1M4 repeat: matches 2333. .2682 of consensus"

repeat_region
repeat_region
repeat_region
repeat_region


```
AX321258      AX321258      395 bp      DNA      linear      PAT 15-DEC-2001
LOCUS          Sequence 275 from Patent W00177168.
ACCESSION      AX321258
VERSION        AX321258.1  GI:17904913
KEYWORDS
SOURCE         human.
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1
AUTHORS        Lodes,M.J., Wang,T., Mohamath,R. and Indirias,C.Y.
TITLE          Compositions and methods for the therapy and diagnosis of lung
JOURNAL        cancer
PATENT: WO 0177168-A 275 18-OCT-2001;
CORIXA CORPORATION (US)
FEATURES       source
               Location/Qualifiers
               1..395
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
BASE COUNT     86 a 138 c 126 g 44 t 1 others
ORIGIN
Query Match    5.3%; Score 89.4; DB 6; Length 395;
Best Local Similarity 98.9%; Pred. No. 3.1e-15;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1608 AGCGGAGCGGAGAGCGGACCCAGAGAGCCCTGAGCAGCCCGCCACCGCCCGCGGCGCT 1667
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 13 AGGGGAGCGGAGAGCGGAGGAGCCCGGAGAGAGCCCTGAGCAGCCCGCCCGCGGCGCT 72
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1668 AGTTACCATCACACCCCGGGAGGAGCGCGCAG 1698
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 73 AGTTACCATCACACCCCGGGAGGAGCGCGCAG 103

RESULT 12
AR117695
LOCUS          AR117695      1481 bp      DNA      linear      PAT 16-MAY-2001
DEFINITION     Sequence 3 from patent US 6140126.
ACCESSION      AR117695
VERSION        AR117695.1  GI:14098601
KEYWORDS
SOURCE         Unknown.
ORGANISM       Unclassified.
REFERENCE      1 (bases 1 to 1481)
AUTHORS        Bennett,C.Frank. and Cowseert,L.M.
TITLE          Antisense modulation of Y-box binding protein 1 expression
JOURNAL        Patent: US 6140126-A 3 31-OCT-2000;
FEATURES       Location/Qualifiers
               source
               1..1481
               /organism="unknown"
BASE COUNT     437 a 381 c 386 g 277 t
ORIGIN
Query Match    5.2%; Score 89; DB 6; Length 1481;
Best Local Similarity 100.0%; Pred. No. 5.8e-15;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1610 CGGAGCGGAGAGCGGAGCGGACCCAGAGAGCCCTGAGCAGCCCGCCCGCGGCGCTAG 1669
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2 CGGAGCGGAGAGCGGAGCGGACCCAGAGAGCCCTGAGCAGCCCGCCCGCGGCGCTAG 61
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1670 TTACCATCACACCCCGGGAGGAGCGCGCAG 1698
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 62 TTACCATCACACCCCGGGAGGAGCGCGCAG 90

RESULT 13
AX409453
LOCUS          AX409453      1481 bp      DNA      linear      PAT 14-JUN-2002
DEFINITION     Sequence 2100 from Patent W00229103.
```

```
AX409453      AX409453.1  GI:21442158
LOCUS          AX409453.1  GI:21442158
KEYWORDS
SOURCE         human.
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1
AUTHORS        Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
TITLE          Gene expression profiles in liver cancer
JOURNAL        Patent: WO 0229103-A 2100 11-APR-2002;
GENE LOGIC INC (US)
FEATURES       Location/Qualifiers
               source
               1..1481
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /note="EMBL/GenBank Accession No. J03827"
BASE COUNT     437 a 381 c 386 g 277 t
ORIGIN
Query Match    5.2%; Score 89; DB 6; Length 1481;
Best Local Similarity 100.0%; Pred. No. 5.8e-15;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1610 CGGAGCGGAGAGCGGAGCGGACCCAGAGAGCCCTGAGCAGCCCGCCCGCGGCGCTAG 1669
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2 CGGAGCGGAGAGCGGAGCGGACCCAGAGAGCCCTGAGCAGCCCGCCCGCGGCGCTAG 61
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1670 TTACCATCACACCCCGGGAGGAGCGCGCAG 1698
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 62 TTACCATCACACCCCGGGAGGAGCGCGCAG 90

RESULT 14
HUMYB1A
LOCUS          HUMYB1A      1481 bp      mRNA      linear      PRI 14-JAN-1995
DEFINITION     Y box binding protein-1 (YB-1) mRNA.
ACCESSION      J03827
VERSION        J03827.1  GI:340418
KEYWORDS        Y box binding protein.
SOURCE         Human lymphoblastoid cell line (Swei) cDNA to mRNA.
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 1481)
AUTHORS        Didier,D.K., Schifffenbauer,J., Woulfe,S.L., Zacheis,M. and
               Schwartz,B.D.
TITLE          Characterization of the cDNA encoding a protein binding to the
               major histocompatibility complex class II Y box
JOURNAL        Proc. Natl. Acad. Sci. U.S.A. 85 (19), 7322-7326 (1988)
MEDLINE        89017190
PUBMED         3174636
COMMENT        Draft entry and clean copy of sequence [1] kindly provided by
               B.D.Schwartz 31-AUG-1988.
FEATURES       Location/Qualifiers
               source
               1..1481
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /map="Unassigned"
               1..1481
               /gene="NFYA"
               127..1080
               /gene="NFYA"
               /note="Y box binding protein-1"
               /codon_start=1
               /protein_id="AAA61308.1"
               /db_xref="GI:340419"
               /db_xref="GDB:G00-128-786"
               /translation="MSSAEETQPPAAPAAPALSAADTKPCTGSGAGSGPGGLTS
               AAPAGDKRVIAIKVLGTGVKFNWNGYGFINKNDTKDFVHQTAIKKNPKVLR
               VGDGTVFDFVVEGEGBEAANYTGPVPGVSGKYAADRNHYRYPRRGPRNYQQ
               NQNSESSEKNEGSAPGQAQRRPYRRPYRMRRPYGRRPQYSNPPVQGEVM
               ECADNQGAGEQGRPVQNNYRCYRFRFRGPPRQRPREDGNEEDKENQDGTQSQQP
```

BASE COUNT 437 a 381 c 386 g 277 t
ORIGIN 2 bp upstream of NCII cut site.
PORRYRRNRYRRRPENPKPODGKNETKRAADPPAENSRSG"

Query Match 5.2%; Score 89; DB 9; Length 1481;
Best Local Similarity 100.0%; Pred. No. 5.8e-15;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1610 CGGAGCGGAGACGCCACCCAGAGACCTGTAGCAGCCCCACCGCGCGCGGCTAG 1669
|||||
Db 2 CGGAGCGGAGACGCCACCCAGAGACCTGTAGCAGCCCCACCGCGCGCGGCTAG 61
|||||

QY 1670 TTACCATCACACCCCGGAGGAGCGGCAG 1698
|||||

Db 62 TTACCATCACACCCCGGAGGAGCGGCAG 90
|||||

RESULT 15

HUMPSDBPB 2168 bp DNA linear PRI 13-OCT-1995
LOCUS Homo sapiens DNA-binding protein B pseudogene (PSDBPB1) gene,
DEFINITION complete cds.

ACCESSION

L37516

VERSION 1.1 GI:950414

KEYWORDS DNA-binding protein B; pseudogene.

SOURCE Homo sapiens

ORGANISM

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

Two human genes isolated by a novel method encode DNA-binding

nucleic-acid-binding proteins containing a cold-shock domain

Gene 73 (2), 499-507 (1988)

JOURNAL

MEDLINE

PUBMED

2977358

REFERENCE

2 (bases 1 to 2168)

Kudo, S., Mattei, M.G. and Fukuda, M.

Characterization of the gene for dbpA, a family member of the

nucleic-acid-binding proteins containing a cold-shock domain

Eur. J. Biochem. 231 (1), 72-82 (1995)

JOURNAL

MEDLINE

PUBMED

7628487

FEATURES

source

1..2168

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="NAR12"

/tissue_type="placenta"

427..1978

/gene="PSDBPB1"

427..445

/evidence-experimental

/rpt_type-direct

581..1540

/gene="PSDBPB1"

/note="cold shock domain; RNP-1 motif"

/pseudo

/codon_start=1

/product="DNA-binding protein B"

1922..1927

/gene="PSDBPB1"

1960..1978

/evidence-experimental

/rpt_type-direct

BASE COUNT 664 a 531 c 510 g 463 t

ORIGIN

Query Match 5.1%; Score 87.4; DB 9; Length 2168;

Best Local Similarity 97.1%; Pred. No. 1.9e-14;

Matches 100; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 1596 AGTTCGATCGGTACGGGAGCGGAGCGGACCCCGAGAGACCTGTAGCAGCCCCACCGC 1655

Db 444 AGTTCGATCGGTACGGGAGCGGAGCGGACCCCGAGAGACCTGTAGCAGCCCCACCGC 503
|||||

QY 1656 CGCCGCCCGCCTAGTTACCATCATCACCCCGGGAGGAGCGCAG 1698
|||

Db 504 CG--GCCGCGCTAGTTACCGTCTACACCCCGGGAGGAGCGCAG 544
|||||

Search completed: November 24, 2002, 18:06:36
Job time : 5110 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2002, 15:01:58 ; Search time 348 Seconds
(without alignments)
10988.199 Million cell updates/sec

Title: HSYBL_COPY_453_2150

Perfect score: 1698
Sequence: 1 AGAACCATATGTCGCCAA.....CACCCCGGAGGAGCCGACG 1698

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

1:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21:	/SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22:	/SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23:	/SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24:	/SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	469.6	27.7	2102	23	AA583003
2	111	6.5	178	21	AAC09730
3	107.4	6.3	1548	20	AA277507
4	89.4	5.3	395	24	AA561724
5	89	5.2	1481	22	AA81319
6	88.6	5.2	1481	22	ABN95602
7	85.2	5.0	179	22	AA02684
8	85.2	5.0	410	23	AA023430
9	83	4.9	3073	12	AA014635
DNA encoding novel					
Human secreted pro					
Human ovarian tumo					
Lung small cell ca					
Human Y-box bindin					
Gene #2100 used to					
Human reproductive					
Human lung tumour-					
Clone associated w					

10	83	4.9	3073	17	AA734371	Plasmid pATG29 (AT
11	83	4.9	3073	20	AAZ32246	Human glioblastoma
12	83	4.9	3073	21	AA88181	pATG29 human gliob
13	81.4	4.8	1474	24	ABK84103	Human cDNA differe
14	78.8	4.6	381	21	AA65939	Human lung cancer-
15	78.8	4.6	381	24	ABL49158	Human lung tumour
16	77.4	4.6	545	24	AA561751	Lung small cell ca
17	61.2	3.6	2179	22	AAH15631	Human cDNA sequenc
18	47.4	2.8	60	24	ABN49420	Human spliced tran
19	45.2	2.7	3900	22	AA526731	Human genomic DNA
20	45.2	2.7	3901	22	AA526729	Human immunoglobul
21	44.8	2.6	3152	22	AA528912	Sheep PrP gene for
22	41.8	2.5	4140	10	AA92735	Mouse ischaemic co
23	41.6	2.4	1469	24	AB199233	Gene encoding a su
24	41.4	2.4	10732	21	AAAI0594	TNFR/NGER protein
25	41.2	2.4	2452	24	AAI71445	Human musculoskele
26	41	2.4	602	22	AAI37306	Clone of recombin
27	40.8	2.4	13104	14	AA046852	Human chemically p
28	40.2	2.4	6731	24	ABK39863	Human immune syste
29	39.8	2.3	5195	24	ABL32920	Human DNA for stag
30	39.8	2.3	6944	24	ABK34026	Human ORFX ORF1824
31	38.8	2.3	584	21	AA676269	Human DNA represen
32	38.8	2.3	98690	24	ABK12169	DNA encoding novel
33	38.4	2.3	160	22	AA543317	DNA encoding novel
34	38.4	2.3	327	22	AA543392	Human polynucleoti
35	38.4	2.3	327	22	AAI84918	Human cDNA sequenc
36	38.4	2.3	1952	22	AAH15912	Human cDNA sequenc
37	38.4	2.3	2287	22	AAH17021	Human breast speci
38	38.4	2.3	2399	24	ABT03034	Human single nucle
39	38	2.2	130	20	AAH86086	Human secreted pro
40	38	2.2	167	21	AA621418	Colon adenocarcino
41	38	2.2	413	24	ABL62435	Ovary cancer relat
42	38	2.2	413	24	ABL67984	Human prostate exp
43	37.6	2.2	738	23	ABV23910	Human prostate exp
44	37.6	2.2	738	23	ABV29791	DNA transcription
45	37.2	2.2	6167	24	ABK28370	

ALIGNMENTS

RESULT 1
AAS83003
ID AAS83003 standard; cDNA; 2102 BP.
XX AC AAS83003;
XX AC
DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #18807.
XX Human; chromosome mapping; gene mapping; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-0508631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG18816.
XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
XX biodiversity -

PS Claim 1: SEQ ID No 19807; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 2102 BP; 580 A; 540 C; 547 G; 434 T; 1 other;

Query Match 27.7%; Score 469.6; DB 23; Length 2102;
Best Local Similarity 97.8%; Pred. No. 2.3e-138;
Matches 497; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

QY 837 AGTAAAGAAAGTGTGAATCTTGGAAAGCAGACACATTTTTTAAATATATGCTGGT 896

Db 1 AGTAAAGAAAGTGTGAATCTTGGAAAGCAGACACATTTTTTAAATATATGCTGGT 60

QY 897 AAAACGGATCAGAGCAGGTCCCATCTGGAGCAGACACCTCGCCCTAAACATGCTGAACCC 956

Db 61 AAAACGGATCAGAGCAGGTCCCATCTGGAGCAGACACCTCGCCCTAAACATGCTGAACCC 120

QY 957 GGGCTGCATAGCTGGTGGTCCCTCCAGGTGACTGCTCCGACAAAGGGTACGCTCT 1016

Db 121 GGGCTGCATAGCTGGTGGTCCCTCCAGGTGACTGCTCCGACAAAGGGTACGCTCT 180

QY 1017 TCAACGSCA-TACGTTTAAGGCAATCCAGAAACCCCTCGCTGCGCGCACTACACGGC 1075

Db 181 TCAACGCACTAGTTAAGGCAATCCAGAAACCCCTCGCTGCGCGCACTACACGGC 240

QY 1076 CATTAAAGAAAGCAGCTCTATGCCCGCGGTAAATGTTCTCAGATCACAGGACCGTATT 1135

Db 241 CATTAAAGAAAGCAGCTCTACGCCCGCGGTAAATGTTCTCAGATCACAGGACCGTATT 300

QY 1136 TGGAGCTGGAGGGAGGAGCCCTTTTCTTCACGGGGGGCTAAGGGCTCTTCGAGCCCCC 1195

Db 301 TGGAGCTGGAGGGAGGAGCCCTTTTCTTCACGGGGGGTAAAGGGCTCTTCGAGCCCCC 360

QY 1196 TTCAATCCCGGTTCGGCGGGTAAATCCCTGCCAGCGTTTCGGGGTGCCTTTTTC 1255

Db 361 TTCAATCCCGGTTCGGCGGGTAAATCCCTGCCAGCGTTTCGGGGTGCCTTTTTC 420

QY 1256 GCCGAGACAAACCTGAAGCTGGGGGGCGCCGAGCCCGGGCTGCTCGTGGAGTCA 1315

Db 421 GCCGAGACTCAACCTGAAGCT-GGGGGCGCCGAGCCCGGGCTGCTCGTGGAGTCA 479

QY 1316 CGTCTCTTCGCGGCTCTTCGGGTAC 1343

Db 480 CGTCTCTTCGCGGCTCTTCGGGTAC 507

RESULT 2
AAC09730

ID AAC09730 standard; cDNA; 178 BP.

XX AAC09730;

XX 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 13805.

XX Human: 5' EST: expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
PS Claim 1: SEQ ID 13805; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.

XX Sequence 178 BP; 38 A; 70 C; 52 G; 17 T; 1 other;

Query Match 6.5%; Score 111; DB 21; Length 178;

Best Local Similarity 99.1%; Pred. No. 9.4e-25;

Matches 111; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1587 ATTCTCGCTAGTTCGATCGGTAGCGGAGCGGAGACGCCACCCAGAGCCCTGAGCAG 1646

Db 1 ATTCTCGCTAGTTCGATCGGTAGCGGAGCGGAGACGCCACCCAGAGCCCTGAGCAG 60

QY 1647 CCCACCGCGCGCGCGCTAGTTACCATCACACCCCGGAGGAGCCGCAG 1698

Db 61 CCCACCGCGCGCGCGCTAGTTACCATCACACCCCGGAGGAGCCGCAG 112

RESULT 3

ID AAZ77507 standard; cDNA; 1548 BP.

XX AAZ77507;

XX 10-APR-2000 (first entry)

XX Human ovarian tumor cDNA library derived EST fragment 58.

XX

KW viral gene expression; extracellular matrix degradation regulator;
KW redox signalling; expression inhibition; antisense therapy;
KW tumour formation; cancer multidrug resistance; inflammation;
KW immune disorder; infection; ds.
XX
XX Homo sapiens.
FH Key Location/Qualifiers
FT CDS 127..1080
FT FT /*tag= a
FT FT /product= "Human Y-box binding protein"
FT FT /function= "Transcription factor"
FT FT misc_feature 1032..1077
FT FT /*tag= b
FT FT /note= "The 15 amino acids encoded by this
FT FT region are not included in the
FT FT corresponding protein sequence"
XX
XX
XX US6140126-A.
XX
XX 31-OCT-2000.
XX
XX 26-OCT-1999; 99US-0429323.
XX
XX 26-OCT-1999; 99US-0429323.
XX
XX (ISIS-) ISIS PHARM INC.
XX
XX Bennett CF, Cowseart LM;
XX
XX WPI: 2001-023284/03.
XX P-PSDB; AAB29744.
XX
XX Antisense oligonucleotides, useful for modulating the expression of
XX Y-box binding protein 1, as well as for treating or preventing diseases
XX associated with Y-box binding protein 1 expression, e.g. inflammation
XX or tumor formation
XX
XX Example 15; Column 47-50 ; 40pp; English.
XX
XX This sequence represents DNA encoding human Y-box binding protein 1.
XX Human Y-box binding protein 1 (also known as YB-1, DNA binding protein B
XX and dbpB) is a member of the Y-box binding protein family of
XX transcription factors, a highly conserved family of nucleic acid binding
XX proteins which bind to the Y-box, an inverted CCAAT sequence found in
XX the promoters of many genes. Y box binding proteins have a broad
XX specificity for nucleic acids, being able to bind double-stranded DNA,
XX damaged DNA, and single-stranded DNA and RNA. Y-box binding protein 1
XX plays a role in DNA repair and the sensitisation of cells from a diverse
XX array of genotoxic stresses, including DNA cross-linking agents and
XX ultraviolet irradiation. Y-box binding protein 1 is also involved in
XX immune regulation, being a negative regulator of MHC (major
XX histocompatibility complex) gene expression, and additionally modulates
XX viral gene expression. It also participates in the regulation of
XX extracellular matrix degradation, and is thought to be involved in redox
XX signalling. The invention relates to antisense oligonucleotides targeted
XX to the human Y-box binding protein 1 gene, which inhibit its expression.
XX A series of oligonucleotides (AAC81326-C81405) were designed to target
XX different regions of the human Y-box binding protein 1 mRNA, and were
XX analysed for their effect on Y-box binding protein 1 mRNA levels by
XX quantitative real-time PCR. The oligonucleotides of the invention are
XX useful for diagnosis, prevention and treatment of conditions associated
XX with Y-box binding protein 1 expression, such as tumour formation,
XX cancer multidrug resistance, inflammation, immune disorders and certain
XX infections.
SQ Sequence 1481 BP; 437 A; 381 C; 386 G; 277 T; 0 other;
Query Match 5.2%; Score 89; DB 22; Length 1481;
Best Local Similarity 100.0%; Pred. No. 3.8e-17;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1610 CGGAGCGGAGCGGACCCAGAGAGCCCTGAGAGCCCGCCGCGCGGCGCTAG 1669

Db 2 CGGAGCGGAGCGGACCCAGAGAGCCCTGAGAGCCCGCCGCGCGGCGCTAG 61
QY 1670 TTACCATCATCACCCCGGAGGAGCGCGAG 1698
Db 62 TTACCATCATCACCCCGGAGGAGCGCGAG 90
RESULT 6
ABN95602
ID ABN95602 standard; DNA; 1481 BP.
XX
XX AC ABN95602;
XX
XX 13-AUG-2002 (first entry)
XX
XX DE Gene #2100 used to diagnose liver cancer.
XX
XX KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
XX KW metastatic liver tumour; cytostatic; expression profile; disease state;
XX KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
XX OS Homo sapiens.
XX
XX PN WO200229103-A2.
XX
XX PD 11-APR-2002.
XX
XX PF 02-OCT-2001; 2001WO-US30589.
XX
XX PR 02-OCT-2000; 2000US-237054P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX
XX WPI: 2002-426119/45.
XX
XX PT Diagnosing and detecting the progression of liver cancer.
XX PT hepatocellular carcinoma or metastatic liver tumor in a patient.
XX PT involves detecting the level of expression of two or more genes in a
XX liver tissue sample
XX
XX PS Claim 1; SEQ ID NO 2100; 298pp; English.
XX
XX CC The invention relates to a novel method for diagnosing and detecting the
XX CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
XX CC tumour in a patient, and differentiating metastatic liver cancer from
XX CC hepatocellular carcinoma in a patient, involving detecting the level of
XX CC expression of two or more genes represented in ABN93503-ABN97455 in a
XX CC tissue sample. The method of the invention has hepatotropic, and
XX CC cytostatic activity. The method is useful for diagnosing and detecting
XX CC the progression of liver cancer, hepatocellular carcinoma and metastatic
XX CC liver carcinoma in a patient. The method is useful for identifying
XX CC expression profiles which serve as useful diagnostic markers as well as
XX CC markers that can be used to monitor disease states, disease progression,
XX CC drug toxicity, drug efficacy and drug metabolism.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 1481 BP; 437 A; 381 C; 386 G; 277 T; 0 other;
Query Match 5.2%; Score 89; DB 24; Length 1481;
Best Local Similarity 100.0%; Pred. No. 3.8e-17;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1610 CGGAGCGGAGCGGACCCAGAGAGCCCTGAGAGCCCGCCGCGGCGCTAG 1669
Db 2 CGGAGCGGAGCGGACCCAGAGAGCCCTGAGAGCCCGCCGCGGCGCTAG 61
QY 1670 TTACCATCATCACCCCGGAGGAGCGCGAG 1698

Db 62 TTACCATCACACCCGGGAGGAGCGGCAG 90

RESULT 7
AAL02684
ID AAL02684 standard; cDNA; 179 BP.
XX
XX
AC AAL02684;
XX
XX
DT 21-NOV-2001 (first entry)
XX
XX
DE Human reproductive system related antigen cDNA SEQ ID NO: 2685.
XX
XX
KW Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; ss.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200155320-A2.
XX
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01339.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225477.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249246.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.

PT and cAMP, that modify, complement or suppress genetic defect and
 PT correct associated phenotypic alteration
 XX
 PS Disclosure; Page 111; 169pp; English.
 XX
 CC In the specification this sequence is given the SEQ ID NO. 30.
 CC There is no reference to SEQ ID NO.30 in the text. Other sequences
 CC in the specification were isolated from human glioblastoma cells
 CC and encode cyclic nucleotide PDEs and RAS-related polypeptides.
 CC (They were isolated by their ability to complement or suppress
 CC genetic defects in a biochemical pathway involving cAMP or which is
 CC controlled by a RAS protein).
 XX
 SQ Sequence 3073 BP; 897 A; 684 C; 635 G; 857 T; 0 other;
 Query Match 4.9%; Score 83; DB 12; Length 3073;
 Best Local Similarity 100.0%; Pred. No. 5e-15;
 Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1616 CGGAGAGCGGACCCAGAGAGCCCTGAGCAGCCCGCCGCGCGCTAGTTACCA 1675
 Db 19 CGGAGAGCGGACCCAGAGAGCCCTGAGCAGCCCGCCGCGCGCTAGTTACCA 78
 QY 1676 TCACACCCCGGAGAGCCGCGAG 1698
 Db 79 TCACACCCCGGAGAGCCGCGAG 101
 RESULT 10
 AAT34371
 ID AAT34371 standard; cDNA; 3073 BP.
 XX
 AC AAT34371;
 XX
 DT 09-OCT-1996 (first entry)
 XX
 DE Plasmid pATG29 (ATCC 68591) insert.
 XX
 KW Human; glioblastoma; complementation; S. cerevisiae; S. pombe;
 KW Clone S46; strain RS60.15B; RAS2; RAS2(vall19ala15); Xenopus laevis;
 KW S6 protein kinase; Plasmid pML5; strain SRN37; CAP; adenylyl cyclase;
 KW diploid; strain SP365; ras1; ras1::LEU2; mutation; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 2..1112
 FT /*tag= a
 FT /product= RAS-related protein
 XX
 US5527896-A.
 XX
 PN 18-JUN-1996.
 XX
 PD 20-APR-1990; 90US-0511715.
 XX
 PF 19-APR-1991; 91US-0688352.
 XX
 PR 20-APR-1990; 90US-0511715.
 XX
 PA (COLD-) COLD SPRING HARBOR LAB.
 XX
 XX Colicelli JJ, Wigler MH;
 XX WPI; 1996-299902/30.
 XX P-PSDB; AAW00089.
 DR
 DR DNA mols. isolated from human glioblastoma cells - encode
 XX RAS-related or cyclic nucleotide phosphodiesterase proteins
 XX
 PS Claim 1; Column 107-112; 101pp; English.
 XX
 CC The sequences given in AAT34369-71 and AAT34391 represent plasmid
 CC fragments which contain human glioblastoma cell cDNA inserts which do not

CC encode phosphodiesterases. These cDNA's were obtained by complementation
 CC of two genetically altered S. cerevisiae and S. pombe strains. Clone
 CC S46 was selected by complementation in S. cerevisiae strain RS60.15B.
 CC This strain contains a mutant allele of RAS2, RAS2(vall19ala15), which
 CC renders cells unable to grow at 36 deg. C, because such cells are
 CC defective in RAS function at elevated temperatures. Human cDNA's from a
 CC human glioblastoma cell library were selected that could complement this
 CC defect. The deduced sequence of S46 is homologous to a Xenopus laevis
 CC gene that encodes a known protein kinase, the S6 protein kinase.
 CC Plasmid pML5 was selected by complementation in S. cerevisiae strain
 CC SRN37. This strain contains a disrupted allele of CAP, cap::HIS3. CAP
 CC encodes an adenylyl cyclase associated protein of undetermined function.
 CC As a consequence of this gene disruption, SRN37 fails to grow in medium
 CC rich in amino acids. Human cDNA's were chosen which could complement
 CC this defect. Plasmids pATG16 and pATG29 were selected by
 CC complementation in the S. pombe diploid strain SP565. This strain is
 CC homozygous for disruptions of ras1 (ras1::LEU2). As a result of this
 CC mutation, this strain fails to sporulate, and human cDNA's were selected
 CC which could complement this defect. These genes have unknown function.
 XX
 SQ Sequence 3073 BP; 897 A; 684 C; 635 G; 857 T; 0 other;
 Query Match 4.9%; Score 83; DB 17; Length 3073;
 Best Local Similarity 100.0%; Pred. No. 5e-15;
 Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1616 CGGAGAGCGGACCCAGAGAGCCCTGAGCAGCCCGCCGCGCGCTAGTTACCA 1675
 Db 19 CGGAGAGCGGACCCAGAGAGCCCTGAGCAGCCCGCCGCGCGCTAGTTACCA 78
 QY 1676 TCACACCCCGGAGAGCCGCGAG 1698
 Db 79 TCACACCCCGGAGAGCCGCGAG 101
 RESULT 11
 AAZ32246
 ID AAZ32246 standard; cDNA; 3073 BP.
 XX
 AC AAZ32246;
 XX
 DT 19-JAN-2000 (first entry)
 XX
 DE Human glioblastoma cell RAS-related pATG29 encoding cDNA.
 XX
 KW Phosphodiesterase; dunce-like phosphodiesterase; PDE; DPD; cAMP;
 KW RAS-related protein; immunoreactive; detection; genetic defect;
 KW bronchodilation; increased myocardial contractility;
 KW anti-inflammation; ss.
 XX
 OS Homo sapiens.
 XX
 PN US5977305-A.
 XX
 PD 02-NOV-1999.
 XX
 PF 07-JUN-1995; 95US-0474379.
 XX
 PR 01-MAR-1994; 94US-0206188.
 XX
 PR 20-APR-1990; 90US-0511715.
 XX
 PR 19-APR-1991; 91US-0688352.
 XX
 PA (COLD-) COLD SPRING HARBOR LAB.
 XX
 XX Colicelli JJ, Wigler MH;
 XX WPI; 1999-619709/53.
 XX P-PSDB; AAY49814.
 DR
 DR New isolated RAS-related polypeptides and mammalian cyclic nucleotide
 XX phosphodiesterases, used for screening for agents which can modify
 XX complement or suppress genetic defects -
 XX

PS Example 1; Column 115-120; 145pp; English.

XX The present invention describes new isolated RAS-related polypeptides

CC and mammalian cyclic nucleotide phosphodiesterases (PDEs). RAS-related

CC polypeptides are capable of complementing a defective RAS function in

CC yeast. The products can be used for screening for agents which can

CC modify, complement or suppress a genetic defect in a biochemical

CC pathway in which cAMP participates, or in a biochemical pathway which

CC is controlled, directly or indirectly, by a RAS protein and other

CC proteins affecting cell growth and maintenance. Developing agents that

CC will selectively act upon PDEs is directed toward reproducing the

CC desirable effects of cyclic nucleotides, e.g. bronchodilation,

CC increased myocardial contractility, anti-inflammation, yet without

CC causing the undesirable effects, e.g. increased heart rate or enhanced

CC lipolysis. The products can also be used for therapeutic, diagnostic

CC and prognostic uses. AA23229 to AA23285, and AA49803 to AA49830,

CC represent sequences used in the exemplification of the present

CC invention.

XX SQ Sequence 3073 BP; 897 A; 684 C; 635 G; 857 T; 0 other;

Query Match 4.9%; Score 83; DB 20; Length 3073;

Best Local Similarity 100.0%; Pred. No. 5e-15;

Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1616 CGGAGAGCGGAGCCCTGAGCAGCCCGCCACCGCCCGCGCGCTAGTTACCA 1675

Db 19 CGGAGAGCGGAGCCCGCCACCGCCCGCCCGCGCGCTAGTTACCA 78

OY 1676 TCACACCCCGGAGGCGCGCAG 1698

Db 79 TCACACCCCGGAGGCGCGCAG 101

RESULT 12

AAA88181

ID AAA88181 standard; cDNA: 3073 BP.

XX AC AAA88181;

XX 14-DEC-2000 (first entry)

DT PATG29 human glioblastoma cell insert nucleotide sequence SEQ ID NO:31.

DE Detection; mammalian gene; yeast: microorganism; identification;

KW phenotype; characteristic; dunce-like phosphodiesterase; PDE; RAS;

KW RAS-related protein; genetic defect; hybridisation; probe; ss.

XX Homo sapiens.

OS Schizosaccharomyces pombe.

XX US6100025-A.

XX 08-AUG-2000.

XX 01-MAR-1994; 94US-0206188.

XX 20-APR-1990; 90US-0511715.

PR 19-APR-1991; 91US-0688352.

XX (COLD-) COLD SPRING HARBOR LAB.

PA Colicelli JJ, Wigler MH;

PI WPI; 2000-531664/48.

XX P-PSDB; AAB20625.

DR Novel isolated DNA encoding a mammalian cyclic nucleotide

XX phosphodiesterase is present in plasmids pPDE46, pPDE43 or pPDE339 and

PT is used to modify a genetic defect in a biochemical pathway in which

PT cAMP participates

XX Example 1; Column 121-126; 145pp; English.

PS

XX The present invention describes a purified and isolated DNA (I) which

CC encodes a mammalian cyclic nucleotide phosphodiesterase and is an insert

CC present in the plasmids pPDE46 (ATCC 69552), pPDE43 (ATCC 69551) or

CC pPDE339 (ATCC 69550). The DNA molecules are used to modify, complement

CC or suppress a genetic defect in a biochemical pathway in which cAMP

CC participates and are also used as hybridisation probes. The present

CC invention also describes methods for detecting mammalian genes encoding

CC proteins which can function in microorganisms, particularly yeast, to

CC modify, complement, or suppress a genetic defect associated with an

CC identifiable phenotypic alteration or characteristic in the

CC microorganism. AAA88162 to AAA88218 and AAB29614 to AAB20640 represent

CC sequences used in the exemplification of the present invention.

XX SQ Sequence 3073 BP; 897 A; 684 C; 635 G; 857 T; 0 other;

Query Match 4.9%; Score 83; DB 21; Length 3073;

Best Local Similarity 100.0%; Pred. No. 5e-15;

Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1616 CGGAGAGCGGAGCCCTGAGCAGCCCGCCACCGCCCGCGCGCTAGTTACCA 1675

Db 19 CGGAGAGCGGAGCCCGCCACCGCCCGCCCGCGCGCTAGTTACCA 78

OY 1676 TCACACCCCGGAGGCGCGCAG 1698

Db 79 TCACACCCCGGAGGCGCGCAG 101

RESULT 13

ABK84103

ID ABK84103 standard; cDNA: 1474 BP.

XX AC ABK84103;

XX 14-AUG-2002 (first entry)

DT Human cDNA differentially expressed in granulocytic cells #674.

DE Human; ss: granulocytic cell; DNA chip; bacterial infection;

XX viral infection; parasitic infection; protozoal infection;

KW fungal infection; sterile inflammatory disease; psoriasis;

KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;

KW cardiac reperfusion injury; renal reperfusion injury; ARDS;

KW adult respiratory distress syndrome; inflammatory bowel disease;

KW Crohn's disease; ulcerative colitis; periodontal disease;

KW granulocyte activation; chronic inflammation; allergy.

XX Homo sapiens.

OS WO200228999-A2.

XX 11-APR-2002.

XX 03-OCT-2001; 2001WO-US30821.

XX 03-OCT-2000; 2000US-237189P.

PA (GENE-) GENE LOGIC INC.

XX Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;

PI WPI; 2002-435328/46.

XX Detecting granulocyte activation by detecting differential expression

XX of genes associated with granulocyte activation, which serves as

PT diagnostic markers that is useful for monitoring disease states and

PT drug toxicity

XX Claim 1; SEQ ID No 674; 114pp; English.

PS The invention relates to detecting (M1) granulocyte (GC) activation

CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by

CC DNA chip analysis as given in the specification, and comparing
CC the expression level to an expression level in an unactivated
CC CC, where differential expression of Gs is indicative of GCA.
CC Also included are modulating (M2) GA by contacting GC with an agent
CC that alters the expression of at least one gene in Gs; (2) screening (M3)
CC for an agent capable of modulating GCA or an inflammation (especially
CC chronic) in a tissue, an allergic response in a subject, exposure of a
CC subject to a pathogen or sterile inflammatory disease using the
CC gene expression profile; (3) detecting (M4) an inflammation (especially
CC chronic) in a tissue, an allergic response in a subject, exposure of a
CC subject to a pathogen or sterile inflammatory disease, by detecting the
CC level of expression in a sample of the tissue of gene(s) from Gs, where
CC the level of expression of the gene is indicative of inflammation;
CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,
CC an allergic response in a subject, exposure of a subject to a pathogen
CC or sterile inflammatory disease, by contacting a tissue having
CC inflammation with an agent that modulates the expression of gene(s)
CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
CC modulating GA; M3 is useful for screening an agent capable of modulating
CC GCA preferably in an inflammation in a tissue; M4 is useful for
CC detecting an inflammation (especially chronic) in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
CC reperfusion injury, ARDS, adult respiratory distress syndrome,
CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
CC periodontal disease; also bacterial infection, viral infection,
CC parasitic infection, protozoal infection, fungal infection and M5 is
CC useful for treating one of the above conditions. The present
CC sequence represents a gene differentially expressed in granulocytes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 1474 BP; 439 A; 382 C; 388 G; 265 T; 0 other;

Query Match 4.8%; Score 81.4; DB 24; Length 1474;
Best Local Similarity 93.4%; Pred. No. 1e-14;
Matches 85; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1596 AGTTTCGATCGGTACGGGAGCGGAGCGGACCCAGAGAGCCCTGAGCAGCCCGCCACCGC 1655

DB 97 AGTTCAATGCGGTAGCGGAGCGGAGCGTACCCAGAGAGCCCTGGCGACCCACCTC 156

QY 1656 GCGCGCGCGCTAGTTACCATCACACCCCG 1686

DB 157 GCGCGCGCGCTAGTTACCATCACACCCCG 187

RESULT 14

AAC65939

ID AAC65939 standard; cDNA; 381 BP.

XX AAC65939;

XX AAC65939;

XX 21-FEB-2001 (first entry)

XX Human lung cancer-associated cDNA LST-sub6-II-4b.

XX Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;
XX vaccine; detection; ss.

XX Homo sapiens.

XX WO200061612-A2.

XX 19-OCT-2000.

XX 03-APR-2000; 2000WO-US08896.

XX 02-APR-1999; 99US-0285479.

XX 17-DEC-1999; 99US-0466396.

PR 30-DEC-1999; 99US-0476496.

PR 10-JAN-2000; 2000US-0480884.

PR 22-FEB-2000; 2000US-0510376.

XX (CORI-) CORIXA CORP.

XX Wang T, Fan L;

XX WPI; 2000-628399/60.

XX Isolated polypeptide comprising an immunogenic portion of a lung tumor
XX protein is used for detecting and monitoring progression of lung cancer
XX in a patient.

XX Claim 25a; Page 200; 261pp; English.

XX This invention describes a novel isolated polypeptide (I) which
XX comprising an immunogenic portion of a lung tumor protein or variant (P2)
XX which have cytostatic activity. The polypeptides and polynucleotides are
XX used in compositions and vaccines to inhibit the development of cancer,
XX especially lung cancer, in a patient. Methods described in the invention
XX can be used to monitor the progression of a cancer by carrying out the
XX detection at subsequent time points and comparing the results from the
XX different time points. CD4+ and/or CD8+ T-Cells isolated from a patient
XX are treated with P2, polynucleotides encoding P2 or antigen presenting
XX cells expressing P2, and then administered to the patient to inhibit
XX development of cancer.

XX SQ Sequence 381 BP; 83 A; 135 C; 117 G; 41 T; 5 other;

Query Match 4.6%; Score 78.8; DB 21; Length 381;

Best Local Similarity 93.0%; Pred. No. 2.8e-14;

Matches 80; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1613 GAGCGGAGCGGACCCAGAGAGCCCTGAGCAGCCCGCCGCGCGCGCTAGTTA 1672

DB 1 GAGCGGAGCGGACCCAGAGCCCTGAGCAGCCCGCCGCGCGCGCTAGTTN 60

QY 1673 CCATCACACCCCGGAGGAGCCCGCAG 1698

DB 61 NCATCACACCCCGGAGGAGCCCGCAG 86

RESULT 15

ABL49158

ID ABL49158 standard; cDNA; 381 BP.

XX ABL49158;

XX 01-MAY-2002 (first entry)

XX Human lung tumour LST-sub6-II-4b cDNA sequence SEQ ID NO:215.

XX Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
XX immune response; ss.

XX Homo sapiens.

XX WO200200174-A2.

XX 03-JAN-2002.

XX 28-JUN-2001; 2001WO-US21065.

XX 28-JUN-2000; 2000US-0606421.

XX 02-AUG-2000; 2000US-0630940.

XX 21-AUG-2000; 2000US-0643597.

XX 15-SEP-2000; 2000US-0662786.

XX 09-OCT-2000; 2000US-0885696.

XX 12-DEC-2000; 2000US-0735705.

XX 07-MAY-2001; 2001US-0850716.

XX (CORI-) CORIXA CORP.

Search completed: November 24, 2002, 16:41:50
Job time : 385 secs

CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/688,352C
FILING DATE: 19910419
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/511,715
FILING DATE: 20-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Borun, Michael F.
REGISTRATION NUMBER: 25447
REFERENCE/DOCKET NUMBER: 27805/30197
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 346-5750
TELEFAX: (312) 984-9740
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 3073 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 3..1111
US-07-688-352C-31

Query Match 4.9%; Score 83; DB 1; Length 3073;
Best Local Similarity 100.0%; Pred. No. 3.2e-16;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1616 CGGAGAGCGGACCCCGAGAGAGCCCTGAGCAGCCCCACCGCGCGCGCTAGTTACCA 1675
|||||
Db 19 CGGAGAGCGGACCCCGAGAGAGCCCTGAGCAGCCCCACCGCGCGCGCTAGTTACCA 78

QY 1676 TCACACCCCGGAGGAGCCCGCAG 1698
|||||
Db 79 TCACACCCCGGAGGAGCCCGCAG 101

RESULT 3
US-08-474-379C-31
Sequence 31, Application US/08474379C
Patent No. 5977305
GENERAL INFORMATION:
APPLICANT: Wigler, Michael H.
APPLICANT: Colicelli, John J.
TITLE OF INVENTION: CLONING BY COMPLEMENTATION AND RELATED
TITLE OF INVENTION: PROCESSES
NUMBER OF SEQUENCES: 88
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,379C

FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/511,715
FILING DATE: 20-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/206,188
FILING DATE: 01-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/688,352
FILING DATE: 19-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 27866/32771
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 3073 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 3..1109
US-08-474-379C-31

Query Match 4.9%; Score 83; DB 2; Length 3073;
Best Local Similarity 100.0%; Pred. No. 3.2e-16;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1616 CGGAGAGCGGACCCCGAGAGAGCCCTGAGCAGCCCCACCGCGCGCGCTAGTTACCA 1675
|||||
Db 19 CGGAGAGCGGACCCCGAGAGAGCCCTGAGCAGCCCCACCGCGCGCGCTAGTTACCA 78

QY 1676 TCACACCCCGGAGGAGCCCGCAG 1698
|||||
Db 79 TCACACCCCGGAGGAGCCCGCAG 101

RESULT 4
US-09-146-249A-31
Sequence 31, Application US/09146249A
Patent No. 6069240
GENERAL INFORMATION:
APPLICANT: Wigler, Michael H.
APPLICANT: Colicelli, John J.
TITLE OF INVENTION: Cloning by Complementation and Related
TITLE OF INVENTION: Processes
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/146,249A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/511,715
FILING DATE: 20-APR-1990
ATTORNEY/AGENT INFORMATION:

```
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3073 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..1111
; US-09-146-249A-31

Query Match
; Sequence 31, Application US/08206188B
; Best Local Similarity 100.0%; Pred. No. 3.2e-16;
; Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1616 CGGAGAGCGGACCCAGAGAGCCCTGAGCAGCCCGCCAGCCCGCGCGGCTAGTTACCA 1675
Db 19 CGGAGAGCGGACCCAGAGAGCCCTGAGCAGCCCGCCAGCCCGCGCGGCTAGTTACCA 78

Qy 1676 TCACACCCCGGAGGAGCGCGAG 1698
Db 79 TCACACCCCGGAGGAGCGCGAG 101

RESULT 5
; Sequence 31, Application US/08206188B
; Patent No. 6100025
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John J.
; TITLE OF INVENTION: Cloning by Complementation and Related
; TITLE OF INVENTION: Processes
; NUMBER OF SEQUENCES: 84
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/206,188B
; FILING DATE: 01-MAR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/511,715
; FILING DATE: 20-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25447
; REFERENCE/DOCKET NUMBER: 27805/30197
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3073 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..1111
; PCT-US91-02714-30

Query Match
; Sequence 30, Application PC/TUS9102714
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John J.
; TITLE OF INVENTION: Cloning by Complementation and Related
; TITLE OF INVENTION: Processes
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/02714
; FILING DATE: 19910419
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/511,715
; FILING DATE: 20-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25447
; REFERENCE/DOCKET NUMBER: 27805/30197
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3073 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..1111
; PCT-US91-02714-30

Query Match
; Sequence 83; Score 83; DB 3; Length 3073;
; Best Local Similarity 100.0%; Pred. No. 3.2e-16;
; Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1616 CGGAGAGCGGACCCAGAGAGCCCTGAGCAGCCCGCCAGCCCGCGGCTAGTTACCA 1675
Db 19 CGGAGAGCGGACCCAGAGAGCCCTGAGCAGCCCGCCAGCCCGCGGCTAGTTACCA 78

Qy 1676 TCACACCCCGGAGGAGCGCGAG 1698
Db 79 TCACACCCCGGAGGAGCGCGAG 101

RESULT 6
; Sequence 30, Application PC/TUS9102714
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John J.
; TITLE OF INVENTION: Cloning by Complementation and Related
; TITLE OF INVENTION: Processes
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/02714
; FILING DATE: 19910419
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/511,715
; FILING DATE: 20-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25447
; REFERENCE/DOCKET NUMBER: 27805/30197
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3073 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..1111
; PCT-US91-02714-30

Query Match
; Sequence 83; Score 83; DB 5; Length 3073;
; Best Local Similarity 100.0%; Pred. No. 3.2e-16;
; Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



```
Db 1408 YYYYYYYYYYYYYYYYYYYYYY 1434
RESULT 9
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOMLOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
; US-08-232-463-14
Query Match 2.6%; Score 44.2; DB 1; Length 7218;
Best Local Similarity 3.8%; Pred. No. 0.002;
Matches 13; Conservative 192; Mismatches 140; Indels 0; Gaps 0;
QY 597 AGAAGGAGAAAGCTGACTCGGAACTATTCATCCCGAGGATAGAGCGCTTAAAA 656
Db 1391 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1332
QY 657 AGAAGGAGAAAGCTGAGTGGGAGGACTTCAGTAACATCAGGTGGCGAGCTCAATTTAT 716
Db 1331 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1272
QY 717 CGTTTGAAACGCTAGTAGTAATCCCTATCATCGGTGGTGTGCAGGAATAAGTAA 776
Db 1271 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1212
QY 777 AAAACAAACAGCTAGCTTCTTCAATAAATGTGAGTTGAATTAATCTGATTGTGTC 836
Db 1211 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1152
QY 837 AGTAGAAAAGATGTGAATCTTGGAAAGGAGACACATTTTTTTAAATATATCGCTGGT 896
Db 1151 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1092
QY 897 AAAACGATCAGAGGAGGAGGTCCCATCGAGACACACCTCGCCT 941
Db 1091 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1047
RESULT 10
US-09-007-005-17/c
; Sequence 17, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihe
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; FILE REFERENCE: 00786/350003
; CURRENT APPLICATION NUMBER: US/09/007,005B
; CURRENT FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(289)
; OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17
Query Match 2.6%; Score 43.4; DB 4; Length 289;
Best Local Similarity 8.2%; Pred. No. 0.00043;
Matches 20; Conservative 104; Mismatches 119; Indels 0; Gaps 0;
QY 223 TGCATTAGCTCTCACTGCTTCTCTACTTCCATCTCTGCGCCCTCTGCAACCATCT 282
Db 243 YGYTYAYCYCYAYCYCYTYGYSYNYNYNYNYNYNYNYNYNYNYNYNYNYNYN 184
QY 283 CCACTCCGCGAGCCATTTTAAAAAGATGCCCTCCCTCTACTTATGACTCTAAATTTGCT 342
Db 183 YSYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYN 124
QY 343 TCTCAGTCTTCCCTCAGGATATATTTCCAATTAATATACCTAAGTACTGCCACCTC 402
Db 123 YSYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYN 64
QY 403 TGAACCCCAATGTACATTCGAGTCTTACTGAGTCTTACTGACTCATTTCCCGAGATCTC 462
Db 63 YSYNYNYCYATYTYGYTYAYATYTYGYTYAYATYTYGYTYAYATYTYGYTYCYC 4
QY 463 ACC 465
Db 3 YCY 1
RESULT 11
US-09-244-796-17/c
; Sequence 17, Application US/09244796
; Patent No. 6281344
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihe
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
```


ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,437
FILING DATE: 05-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DK 88/92
FILING DATE: 23-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: COOPER, Iver P.
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: HANSSON-1A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 13104 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE: ORGANISM: Homo sapiens
IMMEDIATE SOURCE: CLONE: ps459
FEATURE:
NAME/KEY: intron
LOCATION: 1..8834
FEATURE:
NAME/KEY: intron
LOCATION: 8868..10014
FEATURE:
NAME/KEY: intron
LOCATION: 10511..12277
FEATURE:
NAME/KEY: exon
LOCATION: 8835..8867
FEATURE:
NAME/KEY: exon
LOCATION: 10015..10510
FEATURE:
NAME/KEY: exon
LOCATION: 12278..12443
US-08-462-437-4

Query Match 2.4%; Score 40.8; DB 4; Length 13104;
Best Local Similarity 71.1%; Pred. No. 0.038;
Matches 54; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 705 CCTCAATTTTATCTGTTGTAACGTTGATAGTAATCCCTCTATCAGTGGCTGTTGCAG 764

Db 4531 CCTCAGTTTCATCTCTGTAGAATGAGTACTTAATCTCCCTATGTTGGAGTGTGAAA 4472

QY 765 GAATAAAGTGAAGAAA 780

Db 4471 GGATGAGTCAGAAA 4456

RESULT 14
US-09-009-913-1
; Sequence 1, Application US/09009913
; Patent No. 6087485

GENERAL INFORMATION:
APPLICANT: AYS Pharmaceuticals, Inc.
TITLE OF INVENTION: Asthma Related Genes
NUMBER OF SEQUENCES: 339
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bozicevic & Reed, LLP
STREET: 285 Hamilton Ave, Suite 200
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/009,913
FILING DATE: 21-JAN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: SEQ-4P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-327-3231
TELEFAX: 650-327-3231
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 72928 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-09-009-913-1

Query Match 2.0%; Score 34.8; DB 3; Length 72928;
Best Local Similarity 52.8%; Pred. No. 10;
Matches 75; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 706 CTCATTTTATCTGTTGTAACGTTGATAGTAATCCCTCTATCAGTGGCTGTTGCAGG 765

Db 66905 CTAAGTTTTCATCTGGAATGGAGTTAATAACATCTACTGCTGTTGTTAAA 66964

QY 766 AATAAAGTGAAGAAAACAAACAGGCTAGCTGTTCAATAAATGTGAGTTGAATAATCT 825

Db 66965 GATTAATTAACAAAGAAATGTGAAGCACCTGAACAAAGCTTGTGAGTAATAATAGT 67024

QY 826 GATTTGTTGTCAGTAGAAAAAG 847

Db 67025 AATTTGGGAATGAACATCAAG 67046

RESULT 15
US-08-961-527-35/c
; Sequence 35, Application US/08961527
; Patent No. 6420135

GENERAL INFORMATION:

APPLICANT: Charles Kunsch

TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

NUMBER OF SEQUENCES: 391

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19250 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-35

```

[illegible]

Search completed: November 24, 2002, 18:36:06
Job time : 276 secs


```
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY.
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850,716A
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 215
; LENGTH: 381
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (17)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (20)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (60)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (61)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (365)
; OTHER INFORMATION: n=A,T,C or G
; US-09-850-716A-215
```

```
Query Match 4.6%; Score 78.8; DB 10; Length 381;
Best Local Similarity 93.0%; Pred. No. 5.5e-15;
Matches 80; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1613 GAGCGGAGCGGAGCCCTGAGCAGCCCTGAGCAGCCCGCGCGCGGCTAGTTA 1672
|||||
Db 1 GAGCGGAGCGGAGCCCTGAGCAGCCCTGAGCAGCCCGCGCGCGGCTAGTTN 60

QY 1673 CCATCACACCCCGGAGGAGCGCGAG 1698
|||||
Db 61 NCATCACACCCCGGAGGAGCGCGAG 86
```

```
RESULT 6
US-09-897-778-215
; Sequence 215, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 215
; LENGTH: 381
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 17, 20, 60, 61, 365
```

```
; OTHER INFORMATION: n = A,T,C or G
US-09-897-778-215
```

```
Query Match 4.6%; Score 78.8; DB 10; Length 381;
Best Local Similarity 93.0%; Pred. No. 5.5e-15;
Matches 80; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1613 GAGCGGAGCGGAGCCCGGAGGAGCCCTGAGCAGCCCGCGCGCGGCTAGTTA 1672
|||||
Db 1 GAGCGGAGCGGAGCCCTGAGCAGCCCGCGCGCGGCTAGTTN 60

QY 1673 CCATCACACCCCGGAGGAGCGCGAG 1698
|||||
Db 61 NCATCACACCCCGGAGGAGCGCGAG 86
```

```
RESULT 7
US-09-833-790-302
; Sequence 302, Application US/09833790
; Patent No. US20020068288A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tongtong
; APPLICANT: Secrist, Heather
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Indirias, Carol Y.
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.512
; CURRENT APPLICATION NUMBER: US/09/833,790
; CURRENT FILING DATE: 2001-04-11
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 302
; LENGTH: 545
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-833-790-302
```

```
Query Match 4.6%; Score 77.4; DB 10; Length 545;
Best Local Similarity 98.7%; Pred. No. 2e-14;
Matches 78; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1620 GAGCGGAGCCCGGAGGAGCCCTGAGCAGCCCGCGCGCGGCTAGTTACCATCAC 1679
|||||
Db 13 GAGCGGAGCCCGGAGGAGCCCTGAGCAGCCCGCGCGGCTAGTTACCATCAC 72

QY 1680 ACCCGGAGGAGCGCGAG 1698
|||||
Db 73 ACCCGGAGGAGCGCGAG 91
```

```
RESULT 8
US-09-764-864-1705/c
; Sequence 1705, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1705
; LENGTH: 3900
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-764-864-1705

Query Match 2.7%; Score 45.2; DB 10; Length 3900;
```

```

Best Local Similarity   72.0%;   Pred. No. 0.0015;
Matches 59; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 217 GAATAGCTCATTACGCTTCAACTGCTTTCTACTTCCTCATTCCTTGCCCTCTGCAACC 276
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2765 GACTGTACAGTACCTCCCAACTTGCTCTCCCTGTTTCCACTCTTGCCCTCTCTACAATC 2706
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 277 CATTTCTCCACTCCGACGCATT 298
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2705 CATTCTCCACCTGAAGCCAGT 2684

RESULT 9
US-09-764-864-1703
; Sequence 1703, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1703
; LENGTH: 3901
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-864-1703
```

	Query Match	2.7%;	Score 45.2;	DB 10;	Length 3901;
	Best Local Similarity	72.0%;	Pred. No. 0.0015;		
	Matches	59; Conservative	0; Mismatches	23; Indels	0; Gaps
OY	217	GATACTGCATTAGCCCTTAACGTGTTTCTCTACTTCCAATTCCTTGCCCCCTCGCAACC	276		
Dd	1137	GACTGCTACAGTAACTTCGCCAACCTGTGTCCTCCTTTCCACTCTTGCCCTCCTCTACAATC	1196		
OY	277	CATTCTCCAATCCGACGCCATT	298		
Dd	1197	CATTCTCCAGCCCTGAAGCCAGT	1218		

```

RESULT 10
US-09-764-877-3671/c
; Sequence 3671, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3671
; LENGTH: 602.
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3671

```

[illegible]

```

Db      125  CTAATCAAGTAAATAAAAA 105

RESULT 11
US-10-001-887-39/c
; Sequence 39, Application US/100011
; Patent No. US200201554641
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Hervé
; APPLICANT: Cafferkey, Robert
; APPLICANT: Sun, Tongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions
; FILE REFERENCE: DEX-0269
; CURRENT APPLICATION NUMBER: US/11-20
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/249
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: 60/252
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
; LENGTH: 2399
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-001-887-39

```

	Query Match	2.3%	Score 38.4	DB 9	Length 2399
	Best Local Similarity	67.5%	Pred. No. 0.16		
	Matches 54	Conservative 0	Mismatches 26	Indels 0	Gaps 0
QY	703	AGCCTCAATTTATCGTTTGTGAAACGTGGATAGTAAATCCCTCTATCACGTGGCTGTTGC	762		
Db	872	TAATTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT			
QY	763	AACCTGAGTTTCTTGCTGTAAATGGGATAATAATACCTATTTTGAAGCTGTTGA	813		
Db	763	AGGAATAAAGTGAATAAACA	782		
Db	812	AAGGCTTTAAGTGAGTTTAATA	793		

```

RESULT 12
US-09-960-352-2899
; Sequence 2899, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wyatt, John C.
; APPLICANT: Mathalagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 2899
; LENGTH: 280
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 13-LIB3058-009-Q1-K1-D9
US-09-960-352-2899

```

[illegible]

ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC007462.2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.7
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.4
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.3
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.8
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.1
US-09-864-761-9089

Query Match		2.1%	Score 36.2;	DB 10;	Length 562;
Best Local Similarity		46.9%	Pred. No. 0.29;		
Matches 145;		Conservative 0;	Mismatches 163;	Indels 1;	Gaps 1;
QY	352	TCCCTCAGGATATTTCCAAATTAATATACCTAAGTGACTGCCACCTCTGCAACCCA	411		
Db	82	TACCATCTGCTTTTATTTCTGGAAGTTGTCTCAGTATCAGGCTACTGACCTCTGCAA-CCT	140		
QY	412	ATGTCACATTCGAGTCTTACTGAAGTACTTGAGTTCGCGAGATCTCACCTCTTCT	471		
Db	141	ATTTATCATTTGCATACAAAATATTCCTTTCTCTTTCCCTCTCCCTTTCTCCTCC	200		
QY	472	CGCTGTACCTCTGCGGGAAGTCAGCCTCCACCTCTCCCTGCTTCCACTCCCAA	531		
Db	201	CTCCTTCCCTCCCTTCCACCTTCCCTCCCTCCCTCCCTCCCTCCCTTCTCTCA	260		
QY	532	ATACTTCGTGGTTTTCAGCTCTGGAGTATTTACCGTGTGGCTGTTAAATTTCTGCCT	591		
Db	261	CTCCCTTCTTCTTCTCTCTCTTTTATGTTTCTCTTCTTCTTCTTCTGAGT	320		
QY	592	CCATCAGAGGCAAGTACTCGGCAACTATTCATCCCGCGGATAGTAGACGCTT	651		
Db	321	GATACATATCTTCTACTACGTTTTCGCTGCCCACTAACTCCCTGGGAAGAAATCAGAAAAA	380		
QY	652	AAAAAGAA	660		
Db	381	AAAAAAA	389		

Search completed: November 24, 2002, 18:35:49
Job time : 165 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2002, 16:33:38 : Search time 2218 seconds
(without alignments)
12398.529 Million cell updates/sec

Title: HSYBL_COPY_453_2150

Perfect score: 1698

Sequence: 1 AGAACATATGTCGGCCAA.....CACCCGGGAGGAGCGGCAG 1698

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:**

1: em_estba:**

2: em_esthum:**

3: em_estlin:**

4: em_estmu:**

5: em_estov:**

6: em_estpl:**

7: em_estro:**

8: em_hic:**

9: gb_estl:**

10: gb_est2:**

11: gb_hic:**

12: gb_est3:**

13: gb_est4:**

14: gb_est5:**

15: em_estfun:**

16: em_estom:**

17: gb_gss:**

18: em_gss_hum:**

19: em_gss_inv:**

20: em_gss_pln:**

21: em_gss_vrt:**

22: em_gss_fun:**

23: em_gss_mam:**

24: em_gss_mus:**

25: em_gss_other:**

26: em_gss_pro:**

27: em_gss_rod:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	519.4	30.6	532	9 AI659243	AI659243 tul8b06.x
C 2	489.4	28.8	503	13 BI830189	BI830189 603072855
C 3	451.6	26.6	565	12 BF111618	BF111618 7132c03.x
C 4	436	25.7	509	9 AA463628	AA463628 zx98g06.s
C 5	425.4	25.1	438	9 AI650982	AI650982 wa96d06.x
C 6	418.4	24.6	431	9 AI268811	AI268811 qo40g10.x

C 7	410.4	24.2	428	9 AI467820	AI467820 t178b11.x
C 8	405.8	23.9	428	9 AI208001	AI208001 qg60806.x
C 9	399.4	23.5	412	14 N67996	N67996 yz32e04.s1
C 10	354.4	20.9	371	10 AW448961	AW448961 UI-H-B13-
C 11	354.2	20.9	378	12 BE58954	BE58954 7g44f12.x
C 12	352.8	20.8	866	12 BF570994	BF570994 602075915
C 13	332	19.6	434	13 BI828640	BI828640 603078549
C 14	217.6	12.8	238	9 AI640509	AI640509 wa27t07.x
C 15	190	11.2	803	17 A0746664	A0746664 HS_2274.A
C 16	186.8	11.0	1016	11 AK007113	AK007113 Mus muscu
C 17	140.6	8.3	279	13 BI828773	BI828773 603074912
C 18	135	8.0	600	10 BB615022	BB615022 BB615022
C 19	133.4	7.9	217	13 BI550827	BI550827 603195534
C 20	116	6.8	508	9 AU126092	AU126092 AU126092
C 21	114	6.8	1136	13 BM459919	BM459919 AGENCOURT
C 22	114.4	6.7	610	13 BI830400	BI830400 603073451
C 23	112	6.6	403	14 BM853312	BM853312 K-EST0134
C 24	112	6.6	476	14 BM742560	BM742560 K-EST0015
C 25	111.4	6.6	1219	13 BM477621	BM477621 AGENCOURT
C 26	110	6.5	643	10 AW163133	AW163133 au92c01.y
C 27	108	6.4	925	14 BQ421211	BQ421211 AGENCOURT
C 28	108	6.4	960	12 BG387948	BG387948 602412976
C 29	108	6.4	1228	10 BE543630	BE543630 601071354
C 30	106	6.2	950	14 BQ214006	BQ214006 AGENCOURT
C 31	106	6.2	1403	12 BF339614	BF339614 602039089
C 32	105.4	6.2	715	12 BG773027	BG773027 602721301
C 33	104.4	6.1	762	9 AI878912	AI878912 au51c01.y
C 34	103	6.1	165	14 BM847680	BM847680 K-EST0127
C 35	103	6.1	166	14 BM739099	BM739099 K-EST0008
C 36	103	6.1	358	14 BM851331	BM851331 K-EST0132
C 37	103	6.1	360	14 BM850156	BM850156 K-EST0130
C 38	103	6.1	401	14 BM773094	BM773094 K-EST0057
C 39	103	6.1	406	14 BM773061	BM773061 K-EST0057
C 40	103	6.1	413	14 BM845409	BM845409 K-EST0123
C 41	103	6.1	422	14 BM827566	BM827566 K-EST0100
C 42	103	6.1	423	13 BI668317	BI668317 603295680
C 43	103	6.1	437	14 BM784562	BM784562 K-EST0062
C 44	103	6.1	438	14 BM757362	BM757362 K-EST0036
C 45	103	6.1	439	14 BM851642	BM851642 K-EST0132

ALIGNMENTS

RESULT 1

AI659243/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AI659243
tul8b06.x1 NCI-CGAP_Pr28 Homo sapiens cDNA clone IMAGE:2251379 3',
mRNA sequence.
AI659243
AI659243.1 GI:4762813
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 532)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Greg Lennon, Ph.D.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 463.

```
FEATURES
Source
Location/Qualifiers
1..532
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2251379"
/clone_lib="NCI_CGAP_Px28"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: prostate; Vector: pT73D-Pac (Pharmacia)
with a modified polylinker; Plasmid DNA from the
normalized library NCI_CGAP_Px22 was prepared, and ss
circles were used in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (clones IDs
985608-986759, 1101192-1101959, and 1217928-1220615).
Subtraction by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 153 a 92 c 155 g 132 t
ORIGIN
Query Match 30.6%; Score 519.4; DB 9; Length 532;
Best Local Similarity 99.6%; Pred. No. 3.8e-137;
Matches 531; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 245 TCTCTACTTCCATTCCCTGGCCCTCGCAACCCATTCTCCACTCCGCGAGCCATTTTAAA 304
DB 532 TCTCTACTTCCATTCCCTGGCCCTCGCAACCCATTCTCCACTCCGCGAGCCATTTTAAA 473
QY 305 AAGATGCCCTCCCTACTTATGACTCTAAATTTGCTCTCTCACTCTTCCCTCAGGATA 364
DB 472 AAGATGCCCTCCCTACTTATGACTCTAAATTTGCTCTCTCACTCTTCCCTCAGGATA 413
QY 365 TATTTCCAATTAATATACCTAAGTACTGCGCCACCTCTGCAACCCCAATGTCACATTGCA 424
DB 412 TATTTCCAATTAATATACCTAAGTACTGCGCCACCTCTGCAACCCCAATGTCACATTGCA 353
QY 425 GTCTTACTGAATCTACTGACTGATTTCCCGAGATCTCACCTCTTCGCTGTACCTG 484
DB 352 GTCTTACTGAATCTACTGACTGATTTCCCGAGATCTCACCTCTTCGCTGTACCTG 293
QY 485 TGC CGGAAAGTCAAGCCTCCACCTTCCCTGCTTCCACTCCCAAAATACTTCGTGTT 544
DB 292 TGC CGGAAAGTCAAGCCTCCACCTTCCCTGCTTCCACTCCCAAAATACTTCGTGTT 233
QY 545 TTGAGCTCTGAGTATTACCGTGTGGCTGTTTAAATTTCTGCTCCATCAGAAAGCA 604
DB 232 TTGAGCTCTGAGTATTACCGTGTGGCTGTTTAAATTTCTGCTCCATCAGAAAGCA 173
QY 605 GAAACTGACTCGCGAATATTTCATCCCGAGCCGATAGTAGACGCTTAAAGAAAGCA 664
DB 172 GAAACTGACTCGCGAATATTTCATCCCGAGCCGATAGTAGACGCTTAAAGAAAGCA 113
QY 665 AGAAGTGGTGGGAGGACTTCAGTAACATCAGTGGCAGCCTCAATTTTATCGTTTGTG 724
DB 112 AGAAGTGGGT-GGAGGACTTCAGTAACATCAGTGGCAGCCTCAATTTTATCGTTTGTG 54
QY 725 AAACGTGATAGTAATCCCTCTATCAGTGGCTGTTTCAGGAATAAAGTGAAA 777
DB 53 AAACGTGATAGTAATCCCTCTATCAGTGGCTGTTTCAGGAATAAAGTGAAA 1
RESULT 2
BI830189/c
LOCUS BI830189 503 bp mRNA linear EST 04-OCR-2001
DEFINITION 603072855F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5164981 5',
mRNA sequence.
ACCESSION BI830189
VERSION BI830189.1 GI:15941739
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 503)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-femail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11409 row: m column: 14
High quality sequence stop: 499.
Location/Qualifiers
1..503
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5164981"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NOL1;
Site_2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 013. Note:
this is a NIH_MGC Library."

FEATURES
Source

Location/Qualifiers
1..503

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5164981"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NOL1;
Site_2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 013. Note:
this is a NIH_MGC Library."

BASE COUNT 121 a 125 c 117 g 140 t
ORIGIN

Query Match 28.8%; Score 489.4; DB 13; Length 503;
Best Local Similarity 99.8%; Pred. No. 1.4e-128;
Matches 501; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 698 GTGCAGCCTCAATTTTATCGTTTGTGAAAGCTGATAGTAATCCCTCTATCAGCTGGCT 757
DB 502 GTGCAGCCTCAATTTTATCGTTTGTGAAAGCTGATAGTAATCCCTCTATCAGCTGGCT 443
QY 758 GTGCAGCAATAAGTGAACAAACAAACAGGCTAGCTTGTTCATTAATGATGTTGAA 817
DB 442 GTTCAGCAATAAGTGAACAAACAAACAGGCTAGCTTGTTCATTAATGATGTTGAA 384
QY 818 TTAATCTGATTTTGTGCTAGTGAACAAACAGGCTAGCTTGTTCATTAATGATGTTGAA 877
DB 383 TTAATCTGATTTTGTGCTAGTGAACAAACAGGCTAGCTTGTTCATTAATGATGTTGAA 324
QY 878 TTTTAAATATATGCTGTTAAACGGATCAGAGGAGGTCCTCCCATGGAGCACACCTCG 937
DB 323 TTTTAAATATATGCTGTTAAACGGATCAGAGGAGGTCCTCCCATGGAGCACACCTCG 264
QY 938 CCCTAAACATGCTGAACCCGGCTGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 997
DB 263 CCCTAAACATGCTGAACCCGGCTGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 204
QY 998 CGACAAAGGCTAGCTCTTCAACAGCATACGTTTAAAGCAATTCAGAAACCTCGCT 1057
DB 203 CGACAAAGGCTAGCTCTTCAACAGCATACGTTTAAAGCAATTCAGAAACCTCGCT 144
QY 1058 GTGCCGCGACTACAGCCCATTAAGAAAAAGACGACTCTATGCCCCCGTAAATCTTCTCA 1117
DB 143 GTGCCGCGACTACAGCCCATTAAGAAAAAGACGACTCTATGCCCCCGTAAATCTTCTCA 84
QY 1118 GATCAGAGGACCGTATTGAGCTGGAGGGAGGAGCGCTTTTCTTACGGGGGGGCTA 1177
DB 83 GATCAGAGGACCGTATTGAGCTGGAGGGAGGAGCGCTTTTCTTACGGGGGGGCTA 24
QY 1178 AGGCGTCTTCGAGGCCCTTCCA 1200

```

|||||
Db 23 AGCGTCTTCGAGCCCTTCCA 1

RESULT 3
BF111618/c 565 bp mRNA linear EST 20-OCT-2000
LOCUS 7132c03.x1 Soares_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone
DEFINITION IMAGE:3523180 3', mRNA sequence.
ACCESSION BF111618
VERSION BF111618.1 GI:10941308
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 565)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco.
High quality sequence stop: 463.
FEATURES
source
1..565
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3523180"
/clone_lib="Soares_NSF_F8_9W_OT_PA_P_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: p7T3D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not 1; Site_2: Eco RI;
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and ss circles were made in vitro.
Following HAP purification, this DNA was used as tracer in
a subtractive hybridization reaction. The driver was
PCR-amplified cDNAs from pools of 5,000 clones made from
the same 5 libraries. The pools consisted of the following
libraries and cloneIDs: Soares NBHSP pool 1:
309384-310919, 323208-325895 Soares NB2HP pool 1:
145032-147335, 147720-148103, 148872-149255, 15002 -
150407, 151176-152327 Soares NB2HF8-9W pool 1:
758280-760583, 772104-774407 Soares NBHFA pool 1:
304776-306311, 320136-322823, 326280-326563 Soares NBHOT
pool 1: 723720-726407, 739080-740999 Subtraction by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT 151 a 120 c 121 g 173 t
ORIGIN
Query Match 26.6%; Score 451.6; DB 12; Length 565;
Best Local Similarity 98.9%; Pred. No. 9.2e-118;
Matches 465; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Qy 353 CCCTCAGGATATATTTCCAAATTAATATACCTAAGTACGTGCCACCTCTGCAACCCAA 412
Db 472 CACTTGAGGATATATTTCCAAATTAATATACCTAAGTACGTGCCACCTCTGCAACCCAA 413
Qy 413 TGTACATTCGAGTCTACTCACTTACTGACTGCTATTCCTCCGAGATCTCACCTCTCTC 472
Db 412 TGTACATTCGAGTCTACTCACTTACTGACTGCTATTCCTCCGAGATCTCACCTCTCTC 353
Qy 473 GCCTGTACCTGTGCGCGGAAGTCAGCCCTCCACCTTCTCCCTGCTTCCACTCCCAAAA 532
Db 352 GCCTGTACCTGTGCGCGGAAGTCAGCCCTCCACCTTCTCCCTGCTTCCACTCCCAAAA 293
Qy 533 TACTTCGTGGTTTTCAGCTCTGGAGTATTTACCGTGTGGCTGTGTTAAATTTCTGCCTC 592
Db 292 TACTTCGTGGTTTTCAGCTCTGGAGTATTTACCGTGTGGCTGTGTTAAATTTCTGCCTC 233

```

```

|||||
Qy 593 CATCAGAAGCGAGAACTGACTCGCGAACTATTCCATCCCGAGCCGATAGTAGCGTTA 652
Db 232 CATCAGAAGCGAGAACTGACTCGCGAACTATTCCATCCCGAGCCGATAGTAGCGTTA 173
Qy 653 AAAAAGAACGGAAGAGTGGTGGGAGGACTTCAGTAACATCATCAGGTGGCAGCCTCAATT 712
Db 172 AAAAAGAACGGAAGAGTGGTGGGAGGACTTCAGTAACATCATCAGGTGGCAGCCTCAATT 114
Qy 713 TTATCGTTTGTGAAACGCTGGATAGTATCCCTCTATCAGCTGGCTGTGTCAGGGAATAAG 772
Db 113 TTATCGTTTGTGAAACGCTGGATAGTATCCCTCTATCAGCTGGCTGTGTCAGGGAATAAG 54
Qy 773 TGAATAAACAAACACAGGCTAGCTTGTTCATTAATATGAGTTGATTAATAA 822
Db 53 TGAATAAACAAACACAGGCTAGCTTGTTCATTAATATGAGTTGATTAATAA 4

RESULT 4
AA463628/c 509 bp mRNA linear EST 10-JUN-1997
LOCUS zx98g06.s1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:811834
DEFINITION 3', mRNA sequence.
ACCESSION AA463628
VERSION AA463628.1 GI:2188512
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 509)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,
Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wyllie,
T., Waterston, R., and Wilson, R.
WashU-Merck EST Project 1997
Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -41ml3 fwd. Er from Amersham
High quality sequence stop: 486.
FEATURES
source
1..509
/organism="Homo sapiens"
/db_xref="GDB:6042745"
/db_xref="taxon:9606"
/clone="IMAGE:811834"
/clone_lib="Soares_NhHMPu_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: p7T3D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not 1;
Site_2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NBHM, pregnant uterus
NBHPU, and fetal heart NBHH19W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."
BASE COUNT 134 a 109 c 119 g 147 t
ORIGIN
Query Match 25.7%; Score 436; DB 9; Length 509;
Best Local Similarity 98.5%; Pred. No. 2.6e-113;
Matches 461; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

```


QY	353	CCCCCAGGATATTTCCAAATTAATATACCTAAGTGACTGCCACCTCTGCACCCAA	412
Db	467	CACTTGAGGATATTTCCAAATTAATATACCTAAGTGACTGCCACCTCTGCACCCAA	408
QY	413	TGTCACATTCGAGTCTTACTGAACACTACTTGACTGCAATTTCCCGAGATCTCACCTCTCTC	472
Db	407	TGTCACATTCGAGTCTTACTGAACACTACTTGACTGCAATTTCCCGAGATCTCACCTCTCTC	348
QY	473	GCCTGTACCTCTGCGCGGAAGTCAGCCCTCCACCTTCTCCCTGCTTCCACTGCC-AAA	531
Db	347	GCCTGTACCTCTGCGCGGAAGTCAGCCCTCCACCTTCTCCCTGCTTCCACTGCCGAAA	288
QY	532	ATACCTTCGTGGTTTTGTCAGCTCTGGAGTATTTACCGTGTGGCTGTTTTAAATTTCTGCT	591
Db	287	ATACCTTCGTGGTTTTGTCAGCTCTGGAGTATTTACCGTGTGGCTGTTTTAAATTTCTGCT	228
QY	592	CCATCAGAAGCAGAACTGACTCCGGAACATATTCATCCCGCCGCGGATAGACGCTT	651
Db	227	CCATCAGAAGCAGAACTGACTCCGGAACATATTCATCCCGCCGCGGATAGACGCTT	168
QY	652	AAAAAAGAACGGAAGCTGGGTGGGAGGACTTCAGTAACATCAGGTGGCAGCTCAAT	711
Db	167	AAAAAAGAACGGAAGCTGGGTGGGAGGACTTCAGTAACATCAGGTGGCAGCTCAAT	109
QY	712	TTTATCGTTTGAAACCTGGATAGTATATCCCTCTATCACGTGGCTGTTCAGGAATAAA	771
Db	108	TTTATCGTTTGAAACCTGGATAGTATATCCCTCTATCACGTGGCTGTTCAGGAATAAA	49
QY	772	GTGAAAACAAAACAGCTAGCTTGTTCAATAAATGTGAGTTGAATT	819
Db	48	GTGAAAACAAAACAGCTAGCTTGTTCAATAAATGTGAGTTGAATT	1
RESULT	5		
AI650982			
LOCUS	w96d06.x1 NCI_CGAP_GC6	438 bp mRNA linear	EST 16-DEC-1999
DEFINITION	mRNA sequence.		
ACCESSION	AI650982		
VERSION	AI650982.1	GI:4734961	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 438)		
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
COMMENT	Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Bonaldo, Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima		
FEATURES	cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html Insert Length: 552 Std Error: 0.00 Seq primer: -40up from Gibco.		
SOURCE	1. .438		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:2304011"		
	/clone_lib="NCI_CGAP_GC6"		
	/tissue_type="pooled germ cell tumors"		
	/lab_host="DH10B"		
	/note="Vector: pT73D-Pac (Pharmacia) with a modified		

polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA			
from the normalized library NCI_CGAP_GC4 was prepared, and			
ss circles were made in vitro. Following HAP purification,			
this DNA was used as tracer in a subtractive hybridization			
reaction. The driver was PCR-amplified cDNAs from a pool			
of 5,000 clones made from the same library (cloned IDs			
1257096-1258631, 1469084-1470983, and 1475592-1476743).			
Subtraction by Bento Soares and M. Fatima Bonaldo.			
BASE COUNT	101 a	130 c	116 g
ORIGIN			91 t
Query Match	25.1%	Score 425.4;	DB 9; Length 438;
Best Local Similarity	99.8%	Pred. No. 2.6e-110;	
Matches 426;	Conservative	0; Mismatches 1;	Indels 0; Gaps 0;
QY	858	TTGAAAGGAAGACACATTTTTTTTAAATATATGCTGTGTAACGCGATCAGAAGCAGGT	917
Db	3	TTGAAAGGAAGACACATTTTTTTTAAATATATGCTGTGTAACGCGATCAGAAGCAGGT	62
QY	918	CCCCATGGAGCAGACCCCTGCGCCCTAAACATGCTGAACCCGGCTGCCATAGCCTTGGTGG	977
Db	63	CCCCATGGAGCAGACCCCTGCGCCCTAAACATGCTGAACCCGGCTGCCATAGCCTTGGTGG	122
QY	978	TCCCTCAAGGTGACTGCTCCGACAAAAGGTACGCTCTTCAACGCGATACGTTTAAGGC	1037
Db	123	TCCCTCAAGGTGACTGCTCCGACAAAAGGTACGCTCTTCAACGCGATACGTTTAAGGC	182
QY	1038	AATTCCAGAAACCCCTCGGCTGTCCGCGACTACACGGCCATTAAAGAAAAGACGACTCTA	1097
Db	183	AATTCCAGAAACCCCTCGGCTGTCCGCGACTACACGGCCATTAAAGAAAAGACGACTCTA	242
QY	1098	TGCCCGCGGTAAATGTTCTCAGATACAGGACCGTATTGAGCTGGAGGGAGGGAAGC	1157
Db	243	TGCCCGCGGTAAATGTTCTCAGATACAGGACCGTATTGAGCTGGAGGGAGGGAAGC	302
QY	1158	CTTTTCTTCACGGGGGCTAAGGGCTTCGAGCCCGCTTCCATCCCGGCTCGGCCGG	1217
Db	303	CTTTTCTTCACGGGGGCTAAGGGCTTCGAGCCCGCTTCCATCCCGGCTCGGCCGG	362
QY	1218	GTAATCCCTCCAGCGCTTCGGCGTGCCTTTTTCAGCGGACAGACACACCCCTGAACGT	1277
Db	363	GTAATCCCTCCAGCGCTTCGGCGTGCCTTTTTCAGCGGACAGACACACCCCTGAACGT	422
QY	1278	GGGGGCC 1284	
Db	423	GGGGGCC 429	
RESULT	6		
AI268811/C			
LOCUS	q040g10.x1 NCI_CGAP_Lu5	431 bp mRNA linear	EST 17-NOV-1998
DEFINITION	similar to contains OFR.b3 OFR repetitive element.; mRNA sequence.		
ACCESSION	AI268811		
VERSION	AI268811.1	GI:3887978	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 431)		
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
COMMENT	Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D. DNA Sequencing by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be		

found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
 Seq primer: -40UP from Gibco
 High quality sequence stop: 408.

RESULT 7	
AI467820/c	
LOCUS	AI467820 428 bp mRNA linear EST 14-APR-1999
DEFINITION	J178b11.x1 Soares.NSF.F8.9W_OT_PA_P_S1 Homo sapiens cDNA clone IMAGE:2147613 3' similar to contains OFR.b3 OFR repetitive element ;; mRNA sequence.
ACCESSION	AI467820
VERSION	AI467820.1 GI:4329910
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

[illegible]

Best Local Similarity 99.5%; Pred. No. 7e-103;
Matches 411; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

QY 411 AATGTCATTCAGCTTACTGAACACTTGACTGCTATTCGCGAGATCTACCTCTTC 470
      |||
Db 412 AATGTCATTCAGCTTACTGAACACTTGACTGCTATTCGCGAGATCTACCTCTTC 353
      |||
QY 471 TCGCTCTACCTCTGTCGCGGAAAGTCAGCCCTCCACCTTCTCCCTCTTCCACTCCCAA 530
      |||
Db 352 TCGCTCTACCTCTGTCGCGGAAAGTCAGCCCTCCACCTTCTCCCTCTTCCACTCCCAA 293
      |||
QY 531 AATACCTTCGTGGTTTTCAGCTCTGGAGTATTTACCGTGTGGCTGTTTAAATTTCTGCC 590
      |||
Db 292 AATACCTTCGTGGTTTTCAGCTCTGGAGTATTTACCGTGTGGCTGTTTAAATTTCTGCC 233
      |||
QY 591 TCATCAGAGGAGGAGAACTGACTCGGAACTATTTCCATCCCGAGCGATAGACGCT 650
      |||
Db 232 TCATCAGAGGAGGAGAACTGACTCGGAACTATTTCCATCCCGAGCGATAGACGCT 173
      |||
QY 651 TAAAAAGAACGGAAGAGTGGTGGGAGGAGCTTCACTAGTAACATCAGTGGCAGCCTCAA 710
      |||
Db 172 TAAAAAGAACGGAAGAGTGGTGGGAGGAGCTTCACTAGTAACATCAGTGGCAGCCTCAA 114
      |||
QY 711 TTTTATCGTTGTGAACAGTGGATAGTAATCCCTCTATCAGTGGCTGTTCGAGGAATAA 770
      |||
Db 113 TTTTATCGTTGTGAACAGTGGATAGTAATCCCTCTATCAGTGGCTGTTCGAGGAATAA 54
      |||
QY 771 AGTGAAGAAACAAACAGGCTAGCTGTTTCAATAAATGTGAGTTGAATTAAT 823
      |||
Db 53 AGTGAAGAAACAAACAGGCTAGCTGTTTCAATAAATGTGAGTTGAATTAAT 1
      |||

```

```

RESULT 10
AW448961
LOCUS
DEFINITION
  OI-H-B13-ake-a-05-0-01.s1 NCI_CGAP_Sub5 Homo sapiens cdna clone
  IMAGE:2733968 3', mRNA sequence.
ACCESSION
  AW448961
VERSION
  AW448961.1 GI:6989737
KEYWORDS
  EST.
SOURCE
  human.
  ORGANISM
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 371)
  NCI_CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgaaps-r@mail.nih.gov
  The sequence contained an oligo-dT track that was present in the
  oligonucleotide that was used to prime the synthesis of first
  strand cDNA and therefore this may represent a bonafide poly A
  tail. cDNA library Preparation: M.B. Soares Lab Clone distribution:
  NCI_CGAP clone distribution information can be found through the
  I.M.A.G.E. Consortium/LLNL at:
  www-bio.llnl.gov/bbrp/image/image.html
  Seq primer: M13 Forward
  POLYA=Yes.

```

```

FEATURES
  source
    1..371
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:2733968"
    /clone.lib="NCI_CGAP_Sub5"
    /lab_host="DH10B (Life Technologies)"
    /note="Vector: pT73D-Pac (Pharmacia) with a modified
    polylinker; Site:1: Not 1: Site:2: Eco RI; NCI_CGAP_Sub5
    is a subtracted library derived from NCI_CGAP_Sub4. The
    NCI_CGAP_Sub5 library had 3 million recombinants. A
    single-stranded DNA preparation of NCI_CGAP_Sub4 was used
    as a tracer in a subtractive hybridization with a driver

```

comprising: the IMAGE pool (NCI_CGAP_Kid3 pool 1 LLAM 3334-3337, 3682-3683, 3798-3803 (IMAGE Clones 1322376-1323911, 1456008-1456775, 1500552-1502855); NCI_CGAP_Kid5 pool 1 LLAM 3338-3342, 3722-3725, 3776-3778 (IMAGE Clones 1323912-1325831, 1471368-1472903, 1492104-1493255); NCI_CGAP_Lu5 pool 1 LLAM 3575-3582, 3851-3854 (IMAGE Clones 1414920-1417991, 1520904-1522439); NCI_CGAP_GC4 pool 1 LLAM 3164-3167, 3716-3720, 3733-3735 (IMAGE Clones 1257096-1258631, 1460064-1470983, 1475592-1476743); NCI_CGAP_Pr22 pool 1 LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE Clones 985608-986759, 1101192-1101959, 1217928-1220615); NCI_CGAP_Co10 pool 1 LLAM 2644-2653, 2871-2872 (IMAGE Clones 1057416-1061255, 1144584-1145351). (10% of the driver population), plus a pool of 3,840 arrayed clones from NCI_CGAP_Sub1 (IMAGE Clones 2708616-2710535) and NCI_CGAP_Sub2 (IMAGE Clones 2710536-2712455) (10% of the driver population), plus a pool of 11,136 clones from NCI_CGAP_Sub3 (IMAGE Clones 2712456-2723591) (10% of the driver population), plus a pool of 5,472 clones from NCI_CGAP_Sub4 (IMAGE Clones 2723592-2728969) (70% of the driver population). Subtraction was performed as previously described [Bonaldo, Lennon & Soares (1996): Normalization and Subtraction: Two Approaches To Facilitate Gene Discovery. Genome Research 6, 791-806. TAG_LIB=NCI_CGAP_GC4 TAG_TISSUE=germ cell TAG_SEQ=AAATC"

BASE COUNT 89 a 100 c 92 g 90 t
ORIGIN

```

Query Match 20.9%; Score 354.4; DB 10; Length 371;
Best Local Similarity 99.7%; Pred. No. 4.9e-90;
Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 858 TTGGAAGGAGACACATTTTAAATATATATGCTGTAACGATCAGAGGAGGT 917
      |||
Db 16 TTGGAAGGAGACACATTTTAAATATATATGCTGTAACGATCAGAGGAGGT 75
      |||
QY 918 CCCATGGAGACACACCTCCGCCCTAAACATGCTGAACCCGGCTGCATAGCTCGGTGG 977
      |||
Db 76 CCCATGGAGACACACCTCCGCCCTAAACATGCTGAACCCGGCTGCATAGCTCGGTGG 135
      |||
QY 978 TCCCTCAAGTGACTGCTCCGACAAAGGAGTACGCTCTTCAACGATCAGCTTTAAGGC 1037
      |||
Db 136 TCCCTCAAGTGACTGCTCCGACAAAGGAGTACGCTCTTCAACGATCAGCTTTAAGGC 195
      |||
QY 1038 AATTCAGAAACCTCGGCTGTGCGCGACTACACGGCCATTAAGAAAAAGACGACTCTA 1097
      |||
Db 196 AATTCAGAAACCTCGGCTGTGCGCGACTACACGGCCATTAAGAAAAAGACGACTCTA 255
      |||
QY 1098 TGCCGCGGTAAATGTTCTCAGATCAGAGGACCGTATTGAGCTGGAGGAGGAGGAGC 1157
      |||
Db 256 TGCCGCGGTAAATGTTCTCAGATCAGAGGACCGTATTGAGCTGGAGGAGGAGGAGC 315
      |||
QY 1158 CTTTCTTTCACGGGGGCTTAAGGCTCTTCGAGCCCTCTTCAATCCCGGTCCGG 1213
      |||
Db 316 CTTTCTTTCACGGGGGCTTAAGGCTCTTCGAGCCCTCTTCAATCCCGGTCCGG 371
      |||

```

```

RESULT 11
BE858954/c
LOCUS
DEFINITION
  7g44f12.x1 NCI_CGAP_Pr28 Homo sapiens cdna clone IMAGE:3309359 3',
  mRNA sequence.
ACCESSION
  BE858954
VERSION
  BE858954.1 GI:10374531
KEYWORDS
  EST.
SOURCE
  human.
  ORGANISM
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 378)

```

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Greg Lennon, Ph.D.
Cloned Distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov

FEATURES Seq primer: -400p from Gibco
High quality sequence stop: 342.
Location/Qualifiers
1..378
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:309359"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: prostate; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI-CGAP-Pr22 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonoids 985608-986759, 110192-110199, and 1217928-1220615)."
Subtraction by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 100 a 78 c 100 g 100 t

Query Match 20.9%; Score 354.2; DB 12; Length 378;
Best Local Similarity 97.6%; Pred. No. 5.6e-90;
Matches 370; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
QY 404 GCAACCAATGTCACATCGAGTCTTACTGAACTACTGACTGCAATTCCTCCGAGATCTCA 463
DB 378 GCAACCAATGTCACATTCGAGTCTTACTGAACTACTGACTGCAATTCCTCCGAGATCTCA 319
QY 464 CCTCTTCTCGCTGTACCTGTGCGGGAAGTCAGCCCTCCACCTTCCTCCCTGCTTCCA 523
DB 318 CCTCTTCTCGCTGTACCTGTGCGGGAAGTCAGCCCTCCACCTTCCTCCCTGCTTCCA 259
QY 524 CTCCCAAAATACTTCTGCTGTTTGCAGCTCTGAGTATTTACCGTGTGCTGCTTTAAAT 583
DB 258 CTCCCAAAATACTTCTGCTGTTTGCAGCTCTGAGTATTTACCGTGTGCTGCTTTAAAT 199
QY 584 TTCTGCTCCATCAGAACGAGAACTGACTCGGAACTATTCCATCCCGCCGAGTACT 643
DB 198 TTCTGCTCCATCAGAACGAGAACTGACTCGGAACTATTCCATCCCGCCGAGTACT 139
QY 644 AGACGCTTAAAAAGAACGAGAGGTGGTGGGAGGACTTTCAGTAACATCAGTGGCA 703
DB 138 AGACGCTTAAAAAGAACGAGAGGTGGTGGGAGGACTTTCAGTAACATCAGTGGCA 80
QY 704 GCTCAATTTTATCGTTTCTGAAACCTGGATAGTATTCCTCTATCACGTGGCTGTGCA 763
DB 79 GCTCAATTTTATCGTTTCTGAAACCTGGATAGTATTCCTCTATCACGTGGCTGTGCA 20
QY 764 GGAATAAGTGAANAACA 782
DB 19 GGAATAAAAAAAGCCCA 1

RESULT 12
BF570994/c 866 bp mRNA linear EST 12-DEC-2000
LOCUS BF570994

DEFINITION 602075915F1 NIH_MGC_62 Homo sapiens cDNA clone IMAGE:4243371 5',
mRNA sequence.
ACCESSION BF570994
VERSION BF570994.1 GI:11644706
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 866)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC/DCPD/BTP
CDNA Library Preparation: CLONETECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Cloned Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLCM1053 row: m column: 04
High quality sequence stop: 645.

FEATURES Location/Qualifiers
1..866
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4243371"
/clone_lib="NIH_MGC_62"
/tissue_type="melanotic melanoma, high MDR"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggcccgtcggc); Site_2: SfiI (ggccattatggc); SfiI (ggcccgtcggc); Site_2: SfiI (ggccattatggc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-ATTCTAGAGCGGCGGCGCATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, or T). Average insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."
BASE COUNT 240 a 177 c 259 g 189 t 1 others
ORIGIN

Query Match 20.8%; Score 352.8; DB 12; Length 866;
Best Local Similarity 85.8%; Pred. No. 1.18e-89;
Matches 557; Conservative 0; Mismatches 68; Indels 24; Gaps 14;
QY 701 GCAGCCTCAATTTTATCGTTTGTGAAACGTGGATAGTATCCCTCTATCACGTGCTGT 759
DB 638 GCACCTCCAATTTTATCGTTTGTGAAACGTGGATAGTATCCCTCTATCACGTGCTGT 579
QY 760 TCGAGGAATAAGTGAACAAACAGGCTAGCTGTGTTCAATAAAT-GTGAGTTGAAT 818
DB 578 TTCAGGAATAAATGTTTAAACAACAC---GGCTAGCTGTGTTCCATAAATGCTGAGTTGAAT 522
QY 819 TAAATCTGATTTGTGCTCAGTACAAAACATGTAATCTTCGAAAGGA---AGACACAT 875
DB 521 TAAATCTGATTTGTGTGCTCA--TAGACAAGATGTAATCTTGGAAAGGAAAGACACAGCG 464
QY 876 TTTTAAATATATGCTGGTAAACGATCAGAGGC-AGGTCCCTCCATGGAGCACACCC 934
DB 463 TTTGTGAATATATGCTGGTTAAACGGATCAGAGCAAGGTCCCTCCATGGAGCACACCC 404
QY 935 TCGCCCTAAACATGCTGAACCCGGGCTGCCATAGCTGCTGGTGTCTCCCTCAAGG-----T 989
DB 403 TCGCCCT-AACATGCTGGAACCGGGCTTGCATACCTGCTGCTCTCCACAGGTGACT 345
QY 990 GACTGCTCCGACAAAGGCTAGCTCTTCAAACGCTAGCTTTTAAGGCAATTCACCAAC 1049
DB 344 TGTCTCGGTACACAAAGGCTAGCTCTTTTCAACGCTAGCTTTTAAGGCAATTCACCAAC 285

